RELATIONSHIPS AMONG SIX NORTH-EASTERN BANTU LANGUAGES

Magdalena S l a v í k o v á

Being a thesis submitted for the degree of Doctor of Philosophy in the University of London, 1975.

ABSTRACT

Six north-eastern Bantu languages spoken in Kenya are the subject of this thesis. They are Dawida, Saghala, Giryama, Kikuyu, Mvita and Unguja. Saghala, Giryama and Mvita lack substantial primary documentation, Dawida lacks any, and Kikuyu and Unguja are adequately documented.

A study of the kind of relationships these languages display with each other has not been undertaken before, and consequently their place in existing classifications has not been free from ambiguities.

Chapter 1 is a general introduction to the subject of the study, its aims and scope, and the procedures employed.

The comparative approach adopted here is M. Guthrie's methodology as it is presented in his Comparative Bantu: an introduction to the comparative linguistic and prehistory of the Bantu languages (4 volumes, The Gregg Press Ltd., Farnborough, Hants, 1967-1971).

It is based on the examination of lexicons of particular languages and on relating items with a common meaning and regular sound-correspondences. Results of such examination may or may not be interpreted diachronically.

Since this study seeks to establish how close the relationships are between the six selected languages against their common Bantu background, the second chapter contains description of the processes by means of which Common Bantu cognates were identified in each language. It also contains notes on the tendencies, as far as they were found to exist, among sound-correspondences of items which are not perfect cognates of Common Bantu.

Items which have been identified as perfect cognates of Common Bantu are then treated statistically (chapter 3) and several indices are obtained

for degrees of closeness of relationships between each two pairs. The resulting degrees of closeness are then ordered hierarchically and, in conclusion (chapter 4), languages are grouped according to the degree of closeness of relationship between each other, and a tentative diachronic statement is made regarding their likely genealogy.

Table of Contents

ACKNOWLEDGEMENTS

I am indebted to a number of people and institutions who made this study possible.

Grants from the University of London and the School of Oriental and African Studies gave me the opportunity for research, both in London and Kenya.

Professor Malcolm Guthrie, under whose supervision the work started, and on whose initiative the subject of the thesis was chosen, afforded me generously of his time and advice; Professor Wilfred Whiteley, during the brief period of supervision following the death of Professor Guthrie, was a great source of encouragement and enthusiasm to me; Dr. Hazel Carter, who took over after the sudden death of my second supervisor, exerted a great amount of patience and assisted me during the drafting of the actual text. Hers was an unenviable task of taking over work which already had a strong mark of other scholars' influence. I owe her a special thank for discharging her duty with such fortitude. Mr. Michael Mann very kindly undertook the computer work for me. He also made valuable critical comments on the statistics in chapter 3, and, more generally, introduced me to the role of computers in linguistic classification. His unpretentious but firm criticism was greatly appreciated. The present Head of the Africa department in the School of Oriental and African Studies, Professor Edward Ullendorff, very kindly kept stimulating my activity by his interest in my work. For this I am more grateful than he knows.

My informants and friends in Kenya, who patiently provided the very material for this study, may not directly profit by its presentation to the University of London, but use of this work, I do hope, will be made by Kenyan students.

Perhaps the two senior informants, with whom I did most work, can be named on behalf of all. They are the Archdeacon Jeremiah Kiwinda of Shigaro, Wundanyi, who, with his family, helped me in every possible way while I stayed in Wundanyi, and provided an insight not only into the language but into the life-style of the Dawida. Sheikh Yahya Ali Omar of Mombasa spent days with me during his stay at the School of Oriental and African Studies in 1973, discussing the Mombasa dialect of Swahili (Mvita) with me. His interest in the language and his insight were most valuable to me.

Lastly, a word of apology will not be out of place to all those who waited with patience for the conclusion of this work, prodding me along with more (or less) cheerfulness than I deserved.

Magdalena Slavíková,
School of Oriental and African Studies,
University of London,
1975

PREFATORY NOTE

Although I would prefer to sub-title the thesis 'An Introduction', it will, I hope, be the object of further research to expand those areas which will appear fruitful. As a piece of research inspired by Professor Malcolm Guthrie's Comparative Bantu this work intends to make a contribution in an area not covered by him. In this respect it represents but a small initial step in following up Guthrie's pioneering work.

Two interlocking tasks stood at the outset of this study, which were brought to my attention by Professor Guthrie. It was the absence of any appreciable documentation of two languages in the Coast Province of Kenya, namely Dawida and Saghala. Although he included some data from these languages in his Comparative Bantu, they were not among the twenty eight languages which received a detailed attention in that outstanding work. The information it contains on Dawida and Saghala is very limited and certainly invited further research, which he was kind enough to encourage me to do.

The second task was the examination of the relationship between the two languages. It was intended that this relationship would be the focus of the study. But in order to place it in the context of the adjacent languages, Giryama and two Swahili dialects were added to the investigation as representatives of the 'coastal' tendencies which allegedly exist in Dawida and Saghala, and Kikuyu, as a representative of the 'up-country' tendencies which are also to be discerned.

It was a widely but vaguely held view that Dawida and Saghala, while quite closely related, might represent some kind of linguistic watershed between the two influences, the 'coastal' and the 'up-country' one.

Saghala allegedly tended towards the coast, Dawida towards up-country. This kind of feeling was also found to exist among the speakers of these languages, albeit a different stress was placed on it in the two areas. The Saghala feel much more convinced of their alleged 'very close' relationship with the Giryama than the Dawida do about their link with the Kikuyu-Kamba group.

While the relatively close link between Saghala and Giryama has been upheld in this study, the Dawida-Kikuyu link has proved looser than expected. In fact it was found to be looser than Dawida's link with Giryama.

Thus the thesis, while attempting to give answers to some questions, makes it possible for further questions to be posed regarding relationships of the languages in the north-east of the Bantu field.

1 INTRODUCTION

1.1 Choice of subject

Since the attraction of the comparative study of Bantu languages stood at the beginning of this work, a two-tier explanation will be offered here of reasons which led to the choice of this work's two components: the comparative method and the individual languages.

1.1.1 The method

The value of the comparative method as revealing wider aspects of linguistic phenomena than its descriptive counterpart strikes one immediately in African linguistics where the extent to which the method has been employed is still very limited. But the intrinsic value of descriptive linguistics remains appreciated within the statement just made, as indeed it is one of the aims of the present work to make a contribution to Bantu descriptive linguistics by presenting certain results of a primary documentation of several languages.

This author's initial interest in the comparative method in general and in its development by Malcolm Guthrie was essential to the choice of the subject. The publication in 1967 and in consequent years of Guthrie's Comparative Bantu: an introduction to the comparative linguistics and prehistory of the Bantu languages (4 volumes, The Gregg Press Ltd., Farnborough, Hants., 1967-1971) was another inspiration for the undertaking of the research. Prior to the completion of its publication this author had access to some of the unpublished material as well as the benefit of personal discussions with M.Guthrie.

The large-scale examination by Guthrie of more than two hundred Bantu languages provided a new framework within which smaller-scale investigations, such as the present, may be undertaken aiming at different objectives than he did when he set out to make a study of 'all the main aspects of comparative Bantu linguistics /that/ ultimately led to the formulation of a series of hypotheses about Bantu origins.'[1] His work also provides a basis for further comparative investigations which will, it is hoped, shed light on the relationships among the Bantu languages as a whole.[2]

1.1.2 Selection of languages

In this work for the first time Guthrie's comparative method will be applied to original corpora of data, and in a part of the Bantu field not previously subjected to a similar study. The title of the thesis indicates the contribution to the Bantu linguistics made by this study, namely the establishment of affinities among several previously undocumented, or insufficiently documented languages.

A number of factors led to the choice of the six languages which will henceforward be referred to as Selected Languages.[3]

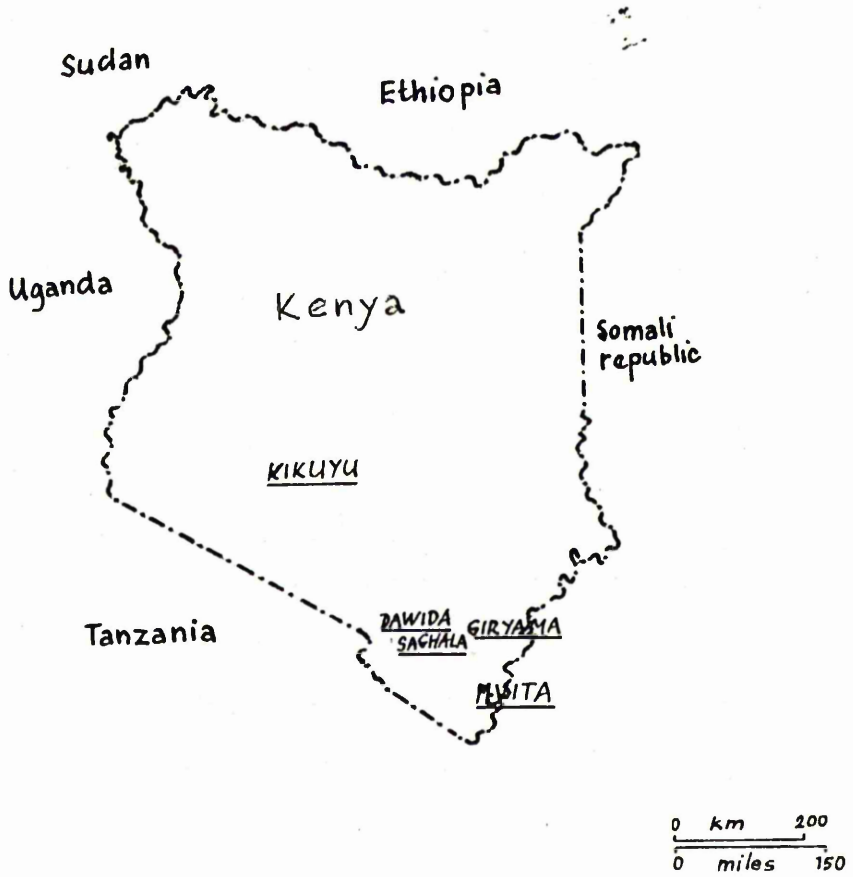
The decisive extra-linguistic factor was this author's knowledge of standard Swahili, a fact which largely determined the geographical area. Thus Swahili could be used throughout the field-work as the linguistic medium. However, the focal point of the research were to be the two

---

[1] M. Guthrie, Comparative Bantu, I, 1, Preface.

[2] The first study of this kind, with reference to Guthrie's comparative Bantu framework, is an unpublished Ph.D. thesis by R. Jones-Phillipson, Affinities between Venda and other Southern Bantu languages, S.O.A.S., University of London, 1972.

[3] In contrast to Guthrie's Test Languages.



Map : Relative geographical positions of the Selected Languages: Kikuyu E.51, Dawida E.74a, Saghala E.74b, Giryama E.72a, Mvita G.42b, and Unguja G.42d.

Taita languages, Dawida and Saghala, which had been largely undocumented, and still lack primary documentation. As Dawida and Saghala appeared to display similarities with the 'coastal' group of languages on the one hand and with the 'up-country' group on the other, representatives were chosen from each group and included in the investigation. Hence Giryama and Kikuyu were incorporated, as well as two Swahili dialects.

Dawida E.74a[1]

The material was initially, and mostly, elicited from the Rev. Jeremiah Kiwinda, about 78, whose speech was representative of the 'Weruga type' dialect of the Dawida language.

Dawida is spoken by a tribe of the same name, who inhabit a mountain massif also called Dawida in the language, and known as the 'Taita Hills' in Kenya. It is situated in the extreme sough-west of the Coast Province of Kenya. There were over 100,000 Dawida speakers in 1970.[2]

All the data used in this work are this author's own although access was made possible to a typewritten copy of an unpublished 'Dawida' vocabulary and a grammar outline.[3]

Since both Dawida and Saghala have been traditionally known in Kenya as 'the Taita language',[4] the fact has been ignored in most classifications that a particular dialect, or local variant, was being reported

---

[1] Classified as E.74a in Guthrie, Classification of Bantu Languages, O.U.P., 1948.

[2] Population figures supplied by the Taita-Taveta District Commissioner and by the District's chiefs.

[3] 'Dawida' in inverted commas, as there was and is not one standard form to be reported. The grammar, and in particular the vocabulary, are marked by their author's strong 'unifying' efforts. Archdeacon Maynard carried out the basic research, while his fellow-missionary Miss M. Murray arranged the material in 1920, and made it available to me.

[4] 'Teita' in German spelling; name taken over from the traditional Swahili terms. 'Taita' refers to the Dawida and Saghala countries, 'Mtaita/Wataita' to the inhabitants (singular and plural), and 'Kitaita' denotes the Dawida and Saghala languages collectively.

under the heading of either Dawida or Saghala, or indeed Taita.[1]

Thus it is important to note that the present work investigates one local variant, or dialect, of Dawida, namely the 'Weruga type'. However, the divergencies within Dawida are not far-reaching, and affect mostly the phonetic structure.

Saghala E.74b

The Saghala material was also collected by this author in the same administrative area of Kenya, namely the Taita-Taveta District of the Coast Province. Altogether six dialectal variants could be distinguished on the Saghala massif near the Voi township and some thirty miles south-east of the Dawida mountains. In 1971 there were some 9,000 speakers of Saghala reported.[2]

The Saghala material presented in this work represents the Teri dialect as it is spoken in the Mlondo area of the Teri Sub-location. Mr. Sebastian, in his forties, and Miss E.A. Matonge, about twenty, were the chief informants.

All data presented throughout this work are this author's own unless otherwise stated.

The Taita Cluster

The one existing monograph on Saghala by J.A. Wray is a telling description of the state of knowledge about the language at that time. Yet this is not to say that the situation has changed radically.

---

[1] J.A. Wray, An elementary introduction to the Taita language, S.P.C.K., London, 1894, is a description of the Teri dialect of Saghala.

[2] Figure supplied by the local chief.

There has been considerable confusion concerning the term 'Taita' linguistically, anthropologically and geographically. More often than not the term denoted 'Dawida', spoken by the people numerically more than ten times stronger and living in an environment far more conducive to a relative economic welfare, while the Saghala were left out of the picture altogether.

The status of the Taita cluster is by no means clear-cut or stabilized in the existing classifications, and it is hoped that the present work will contribute to filling this gap in our knowledge by describing the character of the relationship of the cluster's two components.

Even though Guthrie in his Classification[1] differentiates between the two Taita 'components', he gives them identical dialectal status with that of the 'Nyika dialects' (Giryama, Rabai, Kambe, Ribe, etc.). Notwithstanding Guthrie's degree of generalization the differences between Dawida and Saghala are substantially deeper than those among the Nyika dialects or languages. Not the least among these differences is the absence of two-way interintelligibility of Dawida and Saghala, which exists among the Nyika dialects or languages.

That there is room for further research into the Taita cluster is also apparent from M.A. Bryan's and A.N. Tucker's descriptions of the 'Taita Group' and 'Taita Dialect Cluster'[2] in two different publications. A significant degree of inter-relationships of the Taita cluster and the 'Kikuyu Group' on the one hand and the 'Swahili Group' on the other is apparent from M.A. Bryan's Handbook of the Bantu languages:[3]

---

[1] Classification of the Bantu Languages, O.U.P., 1948.

[2] M.A. Bryan, The Bantu Languages of Africa, O.U.P., 1959; A.N. Tucker and M.A. Bryan, 'Far Eastern Section' in Linguistic survey of the northern Bantu borderland, Vol.1, O.U.P., 1956.

[3] Op. cit.

p.125, Taita Group: Taita

Digo

(but see Nyika in Swahili group)

Pokomo

Nyika

(but see Swahili Group)

*Tharaka, *Chuka

(See also Kikuyu Group)

p.115, Kikuyu Group: Gikuyu

Embu

Meru

*Tharaka

*Chuka

(See also Taita Group)

.

.

.

.

p.126, Swahili Group: Swahili

.

.

.

Nyika : Giryama

.

.

.

Digo

(See also Taita Group)

Note: * indicates 'mixed' languages with affiliations with more than one Group.

Also in Bryan's classification[1] we note the following observation on the 'Swahili Group'[2] and its external relationships.

> 'There is a close relationship between the TAITA and SWAHILI Groups, and definitive classification is not possible without further research.'

The confusion regarding Saghala comes through in the Linguistic Survey of the Northern Bantu Borderland[3] where the Teri dialect is equalled with the Saghala language.

C.M. Doke, too, presents Dawida and Saghala as dialects of a Taita language.[4] In the latest study available on the subject[5] the dialectal status of Dawida and Saghala is replaced by both of them being referred to as 'languages'. Nevertheless, in it, too, the Taita cluster merits a statement that 'Of all the groups we have examined this is the most difficult to judge' [...][6]

---

[1] Op cit., p. 129.

[2] Bryan's 'Swahili Group' consists of all the Swahili dialects together with Nyika dialects or speeches (Bryan's terminology). She also notes that 'The various dialects together known as NYIKA have generally been considered to belong to the TAITA Group, and are thus classified in MG 3.' (p.129.) Following lists of the two groups of dialects, a cross-reference is given to 'the TAITA Group, p.125.'

[3] Op. cit., p.137, and Vol.4, 1957, 63-66.

[4] Doke refers to his classification as a 'tentative grouping [...] based mainly on the arrangement followed by H.H. Johnston in his 'Comparative Study of the Bantu and Semi-Bantu Languages'. See C.M. Doke, Bantu, Dawsons of Pall Mall, London, 1967, p.43.

[5] D. Nurse and G. Philippson, The North-Eastern Bantu Languages of Tanzania and Kenya: a classification, The Institute of Swahili Research, University of Dar es Salaam, 1974.

[6] Op. cit., p.11.

The notion that there are rather close affinities between the Taita cluster and the Kikuyu-Kamba group on the one hand and the 'Mijikenda languages'[1] on the other, has been expressed a number of times.

> 'Certainly the traditions and beliefs of the Wakamba are very similar to those of the Wataita.'[2]

> 'Others [Taita progenitors] came from Ukambani and settled near Ndile; Wagiryama came from Mangea to live on the Sagalla Hill' [...][3]

> 'The Wateita [...], whatever their claims to be called WaNyika, have become to all intents and purposes detached from the body of the Wanyika tribes.'[4,5]

Giryama E.72a

Giryama was chosen to form a 'link' between the Taita languages and Swahili. The Saghala speakers feel a particularly strong affinity with the Giryama and with the other 'Mijikenda languages'. Indeed, they go as far as insisting that theirs is the tenth tribe in the group, while on the other hand they are aware of their manifest affinity to the Dawida.

---

1 Meaning 'Nine-towns languages' in Swahili and often referred to by the Swahili term. A close group of nine rather closely related languages, spoken in vicinity of Giryama.

2 P.G. Bostock, The Peoples of Kenya. The Taita, MacMillan, London, 1950, p.6.

3 Ibid., p.5.

4 A.M. Champion, The Agiryama of Kenya, [ed. J. Middleton], Occasional Paper 25, Royal Anthropological Institute, London, 1967, p.2.

5 'Wanyika' or the 'Nyika tribes' is another commonly used term for the 'Mijikenda'.

Preliminary work had been done before the field-work on W.E. Taylor's Giryama Vocabulary and Collections[1] and Florence Deed's Giryama-English Vocabulary[2] followed by collection in Kenya of verifying data and further additions. Unless otherwise stated, all examples quoted throughout this study are from the present author's own material.

The Giryama 'number over 150,000 and occupy a mainly hinterland area of 2,500 square miles'[3] alongside Kenya coast between Mombasa and Malindi. Yet there is hardly any dialectal variation to be observed within the whole area.

The area around the Kaloleni township, some twenty miles inland from Mombasa, is where the material for this study was assembled. The informants were local residents, both young people as well as several old men. Mr. Christopher Katana, Mr. Gabriel Ngala and Mr. Joseph Gandi, all in their twenties, were the chief informants. They were the only ones among all informants for the Selected Languages who spoke English, but Swahili was always used as the medium with them also.

Kikuyu E.51

Kikuyu was added to the study as possibly the best documented language from the adjacent 'up-country' group to the north-west of Taita. The data used in this analysis are entirely Guthrie's, Kikuyu being one of his twenty eight Test Languages.

All Kikuyu examples quoted in this study were taken from Vols. 3 and 4 of Comparative Bantu.

---

1 Published in 1891.

2 Published in 1964. Taylor's work, however, is more advanced and of a high degree of scholarship. Of particular interest are his phonetic descriptions.

3 D.J. Parkin, Palms, wine and witnesses, Intertext Books, London, 1972, p.7.

Unguja G.42d

This is another of Guthrie's Test Languages, and again, unless otherwise stated, examples quoted will be those from Comparative Bantu.

Guthrie chose as one of his Test Languages the Swahili dialect of Zanzibar town. This dialect forms the basis for Standard Swahili, and it is clear that it is the latter that is being reported. Guthrie later claimed that his own data collected in Zanzibar town in the early 'forties were checked against F. Johnson's Standard Swahili-English Dictionary (O.U.P., London, 1939), which is regarded by Swahili speakers, as well as on comparative evidence, as exhibiting a bias towards the Shambala/Zigula/Zaramo/Bondei group of languages.[1]

For this reason Unguja is paralleled in this study by a comparatively uncontaminated dialect of Swahili, Mvita.

Mvita G.42b

A geographically central dialect of Swahili, Mvita is spoken on Mombasa Island in Kenya. The name is the local term for 'Mombasa'.

The author elicited a Test Language corpus of data for this relatively homogeneous dialect from Sheikh Yahya Ali Omar, an elderly resident of the Old Town of Mombasa. Thus the Mvita material represents an original collection of data, and all examples quoted in this study will be taken from it, unless otherwise stated.

1.1.3 Summary

The comparative Bantu method stood in the focus of this author's

---

[1] M. Slavíková and M.A. Bryan, 'Comparative Bantu: the case of two Swahili dialects', ALS, XIV, 1973, 53-81.

interest when the subject of the present study was being chosen.

The geographical area of the study was largely determined by the author's familiarity with Swahili, which was employed as the linguistic medium throughout the fieldwork in Kenya.

Of the six Selected Languages two were used by Guthrie as Test Languages, whereas the data for the remaining four are the present writer's own, and constitute an original contribution.

For purposes of this study all six corpora of data will be considered as representative of languages rather than dialects.

Nomenclature of languages: The local pronunciation is adhered to, without the independent prefix, as well as the local orthography.[1] The exception is Kikuyu where, for obvious reasons of established traditional orthography, the initial fricative is not represented.

1.2 Aim and scope of work

Having presupposed similarities among the languages selected for this study, the character and extent of these similarities will be identified and described and, eventually, attempts will be made to draw certain diachronic conclusions as to the languages' likely genealogy.

For although the Selected Languages are all manifestly 'north-eastern Bantu', and grouped loosely together in existing classifications, their relationships have not been investigated before. This statement has particular relevance with respect to the Taita cluster and its relationships with the neighbouring Selected Languages.

---

[1] In 'Dawida', the two 'd's' represent implosives, 'w' a bilabial fricative. The initial syllable is stressed. In 'Saghala', 'gh' represents a fricative. Stress penultimate.

1.3 Procedures

The collected data will be juxtaposed and their similarities and differences evaluated first against the background of Common Bantu[1] and then amongst themselves. (Thus comparison is not an exercise for its own sake but a means enabling us to arrive at conclusions about the relationships.)

The main body of investigated data is in the form of a collection of vocabulary items[2] that can be associated in some way with Guthrie's starred forms[3], which symbolize sets of regularly recurrent patterns.

Such patterns, by their presence in the nominal stems and verbal roots, are an essential part of the present investigation of the morpho-phonological structures of the Selected Languages.

In order that this study be comparable both with Comparative Bantu, and with other studies inspired by it, only those methods will be adopted which are presented in it.[4]

---

[1] Guthrie's term referring to the sum total of his Comparative Series (C.S.).

[2] Further data, not directly used in this study, are in the form of tape-recorded conversations, monologues and elicited sentences.

[3] Each Comparative Series (C.S.) is symbolized by a 'starred form'. Guthrie makes further distinction between the starred forms which symbolize sets of regularly recurring correspondences (represented in lower-case letters) and the hypothetical source-items which are inferred from them (represented in upper-case italics).

Throughout the present work 'starred form' is used in reference to the former stage of Guthrie's hypothetical forms. For detailed discussion of the distinction see M. Guthrie, 'A two-stage method in comparative Bantu study', ALS, III, 1962, 1-24.

[4] There have been unusually large numbers reported of Common Bantu reflexes in certain Mijikenda languages. On further scrutiny it was established that Guthrie's criteria of direct reflexes were not strictly adhered to. See W.H. Whiteley, A short list of some C.B. reflexes in Giryama, unpublished typewritten manuscript. A report of a similar research on Ribe was received from W.H. Whiteley in a personal communication.

Although references will be made to CompB throughout this work, it would not serve the purposes of this study if all differences were to be commented on. But where, by using the same procedures significantly different results were obtained, this will be pointed out.

1.3.1 Common Bantu reflexes

Those items that display one-to-one correspondence with their associated starred forms are termed 'direct reflexes'[1] of those starred forms.

It is these direct reflexes that are of main concern in this study: the character and extent of affinities studied here are of languages known prima facie to be genetically related.

Computations have been made in the second stage of this investigation involving direct reflexes, and the resulting figures will be presented and interpreted in terms of the Selected Languages' relationships.

Finally, these relationships will be expounded diachronically.

1.3.2 Skewed reflexes[2]

Certain items which do not display one-to-one correspondence with their associated starred forms are of interest, too, as they may lend weight to arguments about regularity or otherwise of certain patterns. They are an important complement in description of relationships between Bantu languages and Common Bantu.

Since identification of such items raises a number of questions we shall discuss them briefly.

The apparent consequence of identifying skewed reflexes is the bearing it has on increasing or decreasing the number of direct reflexes. Skewed reflexes, although apparently similar to the starred forms in question,

---

[1] Guthrie's term; see CompB, I,1,71.17.

[2] Guthrie's term; see CompB, I,1,32.41-53 and 32.61.

differ from them in some feature. One different feature establishes an item as skewed, but items may be skewed in more than one respect.

a. Semantic and formal skewing.

Items may be skewed for meaning and shape. The former is especially subject to controversy as it is liable to be influenced by subjective judgement. Decisions have to be taken about how remote two synonymous meanings are in order to determine whether an item can be included in a valid Comparative Series or not.

b. Single or multiple, minor or substantial skewing.

Secondly, items may be skewed in more than one aspect. In respect of form, skewings may be single - consonantal, vocalic or tonal - or multiple.

Also, certain skewings may be considered minor, others of a more substantial type. Thus *d>r in place of an expected l or zero would be such a 'minor' skewing, whereas a nasal would be a case of a 'substantial' skewing.

Problems of establishing degrees of divergence in meaning have been mentioned above.

Whereas direct reflexes form a homogeneous body characterized by one-to-one correspondence with respective starred forms both in meaning and shape, the skewed reflexes are far from uniform. Their common characteristic being the lack of correspondence with their starred forms, there is a far-reaching variety among the skewed reflexes by virtue of the generality of that characteristic.

Thus another area of Comparative Bantu, which is outside the scope of this study, is open to further research: a methodology is to be sought for,

which would take into its statistical account not only the direct cognates but also the heterogeneous body of skewed cognates. These latter would subsequently be assorted into types and grades, with special formulae devised in order that they be incorporated in all relevant computations of relationships.

It would follow that Common Bantu, presently formed by direct cognates only, could be extended by a body of 'first degree' skewings and expand considerably the Common Bantu ground of comparative investigation. Such 'first degree' skewings would presumably contain items with a single skewing of 'minor' character first of all, but this approach would demand further investigation into the diffusion of sound-shifts and Guthrie's criterion of regularity. As mentioned previously, such a study of skewed items is beyond the scope of this work which is focussed on relationships among languages as expressed in their direct cognates.

Skewed items have therefore been treated here as a single and, superficially, homogeneous category.

They will be included in the description of Common Bantu reflexes, but not in the computations and eventual diachronic conclusions.

So far a rigorous differentiation and grading of skewings has not figured in comparative Bantu studies[1], but a suggestion has been made for a recognition of 'beside valid and skewed entries, 'potential entries' of unproven validity, a category wanting in Guthrie's presentation.'[2]

---

[1] R.Jones-Phillipson attempted a synthesis of direct and skewed cognates. No 'grading' of skewed cognates was involved, but Guthrie's division into indirect cognates containing no wholly extraneous sounds and those involving wholly extraneous sounds. See 'Affinities between Venda and other Southern Bantu languages', unpublished Ph.D. thesis, S.O.A.S., University of London, 1972.

[2] M. Mann, 'Sound-correspondences and sound-shifts', ALS, XIV, 1973, p.35; a generative approach to the problem of Guthrie's criterion of sound-correspondence regularity.

1.4 Transcription of data

Throughout this work Guthrie's transcription will be followed, such as it is presented in Comparative Bantu.

One important departure from this convention is to be noted: whereas 'r' in Dawida items in CompB represents a trilled lingual 'r' (which occurs in some Dawida dialects), in this work it represents a voiced lateral palatalized fricative 'ɮ' (which occurs in the Weruga-type dialect). This representation was adopted for reasons of economy, and has also been adopted by the local Dawida orthography.

Any further departures from Guthrie's orthographic conventions will be noted in the relevant sections.

LIST OF ABBREVIATIONS AND SYMBOLS

... A	corresponds to the starred form preceded by an augment of the shape indicated
Cl.	concord class
gen	with generalized meaning
intr	intransitive
Mt	metathesis
NA	corresponds to the starred form preceded by *n /homorganic nasal/
ph p	phonaesthetic particle
pl	plural
Rd	reduplicated
sg	singular
spec	with specialized meaning
tr	transitive
YA	corresponds to the starred form preceded by *j
/Y/	corresponds to the starred form with an extension *-j-
C	any consonant
C_1	consonant in first syllable of stem or radical
C_2	consonant in second syllable of stem or radical
C.B.	Common Bantu, sum total of Comparative Series
CompB	Comparative Bantu, by M. Guthrie, Gregg Press Ltd., 1967-71
C.S.	Comparative Series, set of synonymous items from languages in at least three zones displaying regular sound-correspondences
ps.	partial series, as Comparative Series, but confined to two Zones
V	any vowel
V_1	vowel in first syllable of stem or radical
V_2	vowel in second syllable of stem or radical

+	addition to or correction of Guthrie's data
†	multivalent; the item referenced is also suitable for entry under starred form of the shape given in the following note
ǂ	homophonous
Ø	with generalized or specialized meaning. [This is Guthrie's symbol, which this author sometimes replaced by the more specific indications gen or spec.]
Θ	of limited grammatical occurrence
⊖	obsolescent
≠	does not occur in valid entries
*→	the valid shape would be
← *	corresponds to the starred form indicated
>	has as its reflex
⌈ ⌋	skewed shape, *except in abbreviated statements of sound-correspondences, as in 2.5.2, p.94*
( )	skewed meaning, *except in abbreviated statements of sound-correspondences, as in 2.5.2, p.94*

ABBREVIATIONS OF PUBLICATIONS

ALS	African Language Studies, periodical, School of Oriental and African Studies, London.
CompB	Comparative Bantu by M. Guthrie, Gregg Press Ltd., London, 1967-71.
BSOAS	Bulletin of the School of Oriental and African Studies, periodical, London.
JAH	Journal of African History, periodical, Cambridge University Press.

Chapter 2

SELECTED LANGUAGES COMPARED WITH COMMON BANTU

2.1 INTRODUCTION

A broad outline has been made in Chapter 1 of the procedures which will lead first to identification of items which correspond perfectly to their Common Bantu Starred forms and, further on, to computations of direct reflexes in order to establish degrees of relationships among Selected Languages which, in their turn, will lead to certain tentative diachronic conclusions.

At this point it appears necessary to spell out the character and range of the material investigated in the present study.

First of all the total number of 2235 Guthrie's lexical starred forms were employed - together with their individual constituents from various Bantu languages - as a testing screen for the elicited lexical data.[1]

This operation involved not only one-to-one correlation of each lexical item in isolation but its weighting in its morpho-phonemic behaviour. It appears necessary to make this statement here since consideration of morpho-phonemic, or indeed morpho-syntactic behaviour of the studied lexical items still figures in academic polemics.[2]

To state that the lexis only is the subject of the present work would be quite inadequate since in Bantu languages phonology, morphology and syntax

---

1 No doubt there could be more than the 2235 C.S. formed following Guthrie's rules for their construction. But since the existing C.S. provide a sufficient documentation for each Selected Language, and for the sake of comparability with Guthrie's data, the present comparison will be restricted to the existing number of C.S.

2 See J.H.Greenberg, 'Linguistic evidence regarding Bantu origins', JAH, XIII, 2, 1972, 189-190. Greenberg refutes Roland Oliver's alleged suggestion that he employs only lexical evidence while disregarding grammatical evidence.

interlock in determining the shape of lexical items in isolation. It is inherent to CompB methodology that grammatical evidence plays a crucial role in it. Thus e.g., a nominal stem may be accepted as a direct reflex of its starred form on the strength of its either singular or plural form if the two differ.

Thus *-yínò, 'tooth', has as its reflex in Mvita ɟino in Class 5 and meno in Class 6. Since *-yi- disappears without trace in Mvita, the singular form is accepted as a direct reflex whereas its plural, on account of its skewed prefix, is discarded as such. The two Classes, 5 and 6, being treated as a linguistic whole[1], ensure a one point score as a direct reflex to the respective item. In Guthrie's notation the process is recorded as follows:

C.S. 2073 *-yínò 5/6 'tooth' > Mv. ɟino/⌈meno⌋ 5/6.

The same grammatical principle is applied to the treatment of verbal roots, where a cluster of simple and extended verbal roots is treated as one formal unit, too. E.g. *-cànj- 'spread' has two reflexes in Mvita, [-tand-] and -ʄandaz-. The regular sound-correspondence is *c > ʄ and *nj > nd, therefore the causative member of the pair scores one point as containing the regularly derived simplex.[2]

---

1 The term adopted by Guthrie is 'gender'.

2 Guthrie's term for the irreducible core of Bantu verbals, most frequently of the shape -CVC-.

Further examples, taken from Mvita, include:

*-cǎákad- become worn out > /-tʃaka-/, -ʈakalik- spec

*-cǒká 5/6 axe > /ʃoka 5/6/; ki-ʈoka 7/8 spec

It will have been noted throughout the work that individual sound-correspondences have been postulated on bodies of data of different sizes. Thus Swahili k < *k was attested by over ninety items, while correspondences of nasal compounds in close-vowel environment were often attested either by a single item or lacked any evidence at all.

Open-vowel environment provides by definition for many more attesting items than the relatively restricted environment of *$\underset{s}{i}$ and *$\underset{s}{u}$, especially if restricted still further, e.g. as *$\underset{s}{VV}$, in nasal compounds, etc. In such cases supporting evidence was sought either in parallel development of related sound-patterns and/or in structural data.

2.1.1 A note on tone

Although the Selected Languages, Swahili excluded, are tonally distinctive, the entries are not marked. There are several reasons.

Entries in CompB are not tone-marked, which would hinder attempts at correlating tonal data with Kikuyu.

Little attention is paid in CompB to the role of tone in reflexes of C.B., while consideration of the vocalic and consonantal aspects of items appears to prevail.

Although the complex tonal systems of the Taita languages and Giryama would require a separate work to describe them adequately, certain correlation with C.B. tones has been attempted. (All three languages display particular types of tone-shift.) It revealed almost perfect correspondence with C.B. tone marks .

Among the Test-Language body of data for Kikuyu only two (!) items were quoted as tonally skewed.[1]

---

1 Data obtained from the Register of entries from the Test Languages, CompB,I, 2, 65-104 and checked against op.cit., II, 3 and 4.

2.2 REFLEXES OF COMMON BANTU IN DAWIDA

2.2.1 General

Dawida, E.74a being a previously undocumented language it was essential to collect two kinds of vocabulary material for the present work. Firstly, material that would reveal the Dawida structure in its own right, and, secondly, material that would be used for comparison with Common Bantu.

The collected vocabulary comprised approximately 5,000 items not all of which, however, were suitable for the immediate comparative work. As said above, parts of it were used in the prerequisite descriptive stage.

Altogether 303 direct reflexes of Common Bantu starred forms were identified together with extra 127 reflexes that were skewed in some respect.

As Dawida is not one of Guthrie's Test Languages, and therefore only selected items are presented in CompB, it will be useful to identify it in relation to the Test Languages on the scale of the number of C.B. reflexes each language contains. The following table gives the figures, which were supplied by M. Mann. (Dawida E.74a - 303 direct reflexes.)

Duala A.24	215	Lwena K.14	485
Bulu A.74	277	Luba-Lulua L.31	557
Bali A.75	380	Luba-Katanga L.33	606
Bobangi C.32	401	Bemba M.42	800
Tetela C.71	314	Ila M.63	540
Rundi D.62	506	Nyanja N.31	535
Nyoro E.11	446	Yao P.21	579
Nyankore E.13	441	Mbundu R.11	427
Ganda E.15	450	Herero R.31	364
Gikuyu E.51	424	Manyika S.13	508
Kamba E.55	402	Venda S.21	396
Sukuma F.21	547	Sotho S.33	402
Swahili G.42	656 [1]	Xhosa S.41	340
Kongo H.16	618	Zulu S.42	386

Table 1. Common Bantu reflexes in Test Languages.

---

1 The Test Language is the Zanzibar dialect, Unguja G.42d.

2.2.2 The Dawida dialect used

As mentioned in Chapter 1, the material presented in this study is representative of the 'Weruga type' dialect of the Dawida language. It was collected in the Shigaro area of the Taita Hills in south-eastern Kenya. This fact is important to note since with little documented languages different dialects tend to be presented by authors as the one and only existing form of the language. On the other hand, a conglomerate of dialects may be said to be 'the language,' which may give the impression that one particular form of the language is being reported.[1] It is essential to the comparativist to establish the type of dialect, even idiolect, one is dealing with.

Since CompB provided a basis for comparison with this author's data[2] a brief account will be made of the information concerning Dawida, which is contained in it.

First there are the relevant sound-correspondences given[3] and then twenty four Dawida entries are quoted in appropriate C.S. On the strength of that evidence this author is of the opinion that either a dialect different from the 'Weruga type' is reported (most likely the Chawia or Bura variety) or that several sources of information were used,[4]

---

1 Both the New Testament in Dawida and The Book of Common Prayer in the Taita Dabida Language, London, S.P.C.K., 1927 bear signs of containing more than one dialect. So do three consecutive Reading Books in Taita, Sheldon Press, 1938; S.P.C.K. London, 1926; Macmillan, 1957.

2 Broad comparison was also made with the Kidawida Grammar, MS by Archdeacon Maynard and M. Murray, 1920, and an English-Dawida vocabulary, MS, by the same authors.

3 CompB, 1, 2, 47.

4 In a personal communication Guthrie recollected that his data were collected in the Voi township. Voi lies outside the Dawida proper area, at a railway and road junction. It is by far the most 'cosmopolitan' settlement in the whole administrative district. Dawida, Saghala, Swahili and some Taveta, are spoken there.

which would account for instance for the ɣ < *g, which in the 'Chawia-Bura type' would be plosive.[1]

The 'Weruga type' dialect has ɣ as a reflex of *g.

Except in his Test Languages, Guthrie tends to quote only some of the sound-correspondences operating in a language. This is probably the reason why several Dawida correspondences were omitted in CompB, but a different explanation would have to account for the quoted correspondence *t > d, since ɗ < *t does figure in all the dialects.[2] Altogether ten entries are quoted in CompB which contain a d-sound. In nine of them a plain voiced stop is quoted while the tenth entry contains the implosive ɗ.

2.2.3 Direct reflexes of Common Bantu.

A table of Common Bantu-Dawida correspondences is presented below, as it was worked out by the present author. Since not all of the correspondences are equally well documented,[3] some are postulated more tentatively than others.

The following conditions also apply:

1) *VV > V (Common Bantu long vowels correspond to their short equivalents in Dawida)

2) $*C_1 = *C_2$ (sound-correspondences are identical for consonants in first and second positions)

3) *7V > 5V (distinctions *i̱ /*i and *u̱ /*u disappear)

---

1 Cf. the four correspondences taken from CompB. *-dùgad- 'shut,' *-dùgud- 'open' > -rugar-, -rugu-; *-yi̱gut- 'become satiated' > -gud-. But *-gàŋgà 1/2 'medicine man' > m̄ ɣaŋga 1/2; *-gí 5/6 'egg' > i̱ɣi 5/6.

2 Cf. *t > ɗ /[(* ~ i̱) > s/ in CompB, 1, 2, p.47.

3 For the purposes of the present work a 'well documented' sound-correspondence is one which is supported by at least ten occurrences.

Consonant reflexes of Common Bantu in Dawida

(A blank in a column indicates no or inconclusive evidence.)

	*+ -a[1]	*+ -i	*+ -u	*-i+V	*-u+V	*-i+V[2]
*p	ɸ	f	f	f		
*b	β	v	v	v		
*t	ɗ	s	f			
*d	l/ɸ[3]	r[4]	r		r	ʒ
*k	k	s	f		tʃ	tʃ
*g	ɣ	z	v		dʒ	dʒ
*c	s	ʃ		ʃ	ʃ	
*j	tʃ	tʃ	tʃ			
*y	ɸ	ɸ				
*m	m					
*n	n					
*ŋ	ŋ					
*mp	mɓ	f	f			
*mb	mɓ		v			
*nt	nɗ					
*nd	nd					
*ŋk	ŋg	s		ʃ		
*ŋg	ŋg	z				
*nc	s					
*nj	tʃ					

1 *a or any vowel other than high.

2 other than *i.

3 For discussion of the environments in which either 'l' or Zero appear, see the following section 'Consonantal correspondences'.

4 This symbol, for the sake of convenience, represents here a palatalized lateral fricative occurring in the 'Weruga type' of dialect. As it corresponds to a rolled 'r' in all other dialects, the letter 'r' has been used to represent it.

2.2.3.1 Consonant reflexes in open-vowel environment [1]

The regular reflexes of voiceless plosives *p, *t, *k are ɸ, ɗ and k respectively. They are attested by abundant occurrences, and comprise words of common Eastern Bantu stock.

All correspondences presented below will be illustrated by examples taken from this author's data.

*p > ɸ

*-pàká 3/4 boundary > mw-aka 3/4

*-pémbá (3) millet, eleusine > mw-emba 3/4

*-pít- (i) pass (iii)[2] surpass > -iɗ-

*t > ɗ

*-tápik- vomit > -ɗaik-

*tátù DP three > -ɗaɗu

*-tét- speak > -ɗeɗ-

*k > k

*-káán- deny > -kan-

*-kèndá (5)[3] nine > i-kenda 5

*-kúnd- desire > -kund-

The three opposite voiced sounds *b, *d, *g correspond to β, l or ɸ and ɣ respectively. Reflexes of *d will be discussed

---

1 Starred close vowels in CompB are *i̱, *u̱.

2 Roman figures refer to the several applicable glosses as given in CompB. Here, second gloss of the C.S. is not applicable.

3 Number of nominal Class in round brackets means that the Class is attested by less than three items.

last of the three.

*b > β[1]

*-bàbá 5/6, 7/8, 11/10 wing > i-βaβa 5/6

*-bìŋg- chase; chase away > -βiŋg-

*-bón- see > -βon-

*-bùgà 11/10 threshing-floor > ki-βuɣa 7/8

*g > ɣ[2]

*-gàmb- speak > -ɣamb-

*-gègò 5/6 (i) tooth (ii) molar tooth > i-ɣeɣo 5/6

*-gòn- snore > -ɣon-

*-gùbá 3/4 sugar cane > i-ɣuβa 5/6

*d presented a somewhat complex picture.[3] There were altogether 102 reflexes of *d counted among the data. The statistics of the actual reflexes can be summed up as follows:

---

1 'β' is represented by 'w' in the present work in the word 'Dawida' and its derivations, and by 'β' in phonetic transcription of Dawida words. Other conventions existing in older works and in the indigenous area include 'b', 'b̲' and 'ɸ'. So far the Dawida orthography has not been standardized.

2 'ɣ' represented by 'g' or 'gh' in the existing local orthography.

3 However, only *d > r is postulated for open-vowel environment in CompB, 1, 2, p.47.

*d_1 > r ... 4 occurrences[1]	*d_2 > r ... 2 occurrences[4]
*d_1 > l ...32 occurrences[2]	*d_2 > l ...25 occurrences[5]
*d_1 > ∅ ... 7 occurrences[3]	*d_2 > ∅ ...32 occurrences[6]

---

1 *-dàngí 3/4 bamboo > [mu-rangi 3/4]
*-dèngè 5/6 pumpkin > [i-renge 5/6]
*-déét- bring > [-red-]
*-dòngò 3/4, 4 ten > [mu-rongo 3/4 spec]

2 E.g.
*-dámú 1/2 brother-or-sister-in-law > mlamu 1/2
*-dèk- leave tr > -lek-
*-dìngá DP how many? > -linga

3 E.g.
*-dákà voice > lw-aka 11/10a
*-dòndà 7/8 sore > tʃ -onda 7/8
*-dùngò 5/6 joint > β -ungo 14/6

4 *-bàd- split > [-bar-]
*-bèdè 9 front > [m-bere 9]
*-cùùbid- hope > [-subiri-] appears to be a loan from Swahili. In fact the validity of the whole C.S. is doubtful as it contains entries from languages which have been in contact with Swahili; Ganda E.15, Logoli E.41, Saghala E.74b, Nyamwezi F.21, Hehe G.62, Bemba M.42, while the Swahili - Unguja G.42d - entry is a triple skewing, and patently a loan from Arabic صبر subiri 'wait, be patient.'

5 E.g.
*-gàd- come or go back > (-ɣal- go home)
*-gàdud- turn tr > -ɣalu-
*-yéd- shine > -el-
*-yídù DP black > [-ilu]
*-tídó dream > [dilo 9]

6 E.g.
*-gùd- buy > -ɣu-
*-gùdò 5 yesterday > i-ɣuo 5
*-kádá (i) 5/6 ember; charcoal > i-ka/ma-ka 5/6
*-cèdid- come or go down > -se-
*-pùdà 9/10 nose > m-βua 9/10

To the *d > l correspondence was added Class 11 independent prefix lu < *du, e.g. *-gànò 7/8, 11/10 'tale' > lu-ɣano/tʃu-ɣano 11/10a.[1]

Under the $*d_1 > \emptyset$ correspondence was also entered the regular reflex involving Class 7 independent prefix,

e.g. *-dòndà 7/8 > tʃ-onda 7/8

*-dèdù 7/8, 12/13 chin > tʃ-eru 7/8.

The same palatalization process exists in Class 7 nominals with *y in the first consonantal position,

e.g. *-yògà 14 mushroom > tʃ-oɣa 7/8

*-yàmbò 7 bait > tʃ-ambo 7/8

On the strength of the available evidence we shall postulate the following development:

Stage 1: *d > *l

Stage 2: *l > l or ∅

Stage 2 is in fact the existing state, with two optional variants in C_2 position.

---

1 Class 10a has a unique independent prefix tʃu-, while all its dependent affixes are homophonous with corresponding Class 10 affixes. Class 10a forms a regular and productive plural to most Class 11 nominals.
A gender which is formed by Class 11, 10 and 10a is of far less frequency than the 11/10a one. Classes 10 and 10a appear to belu-karangu/tʃu-within this gender. E.g.,
lu-mbala/lu-hehia/tʃu- la 'chip of wood'
Examples of lu-di/tʃu- gender include the following:

lu-finiko/tʃu-	trench
lu-ʃango/tʃu-	lightning
lu-mu/tʃu-	kind of poisonous tree
lu-finiko/tʃu-	lid
lu-ʃango/tʃu-	swamp overgrown with grass
lu-mu/tʃu-	flame

The evidence for C_3 reflexes is rather limited. It is restricted to the directive verbal extension *id and reversive *ud, and the one radical complex *-cèdid-. In the latter case the reflex is zero, while in the former two either zero or r appear. The zero reflex by far predominates.

To conclude, on the strength of the available evidence entries containing an r will be considered skewed, while those containing the optional variants l or zero will be treated as direct reflexes.

*c > s

is another sound-correspondence with sufficient evidence for all consonantal positions, as in *-cábí (i) 1/2 witch (ii) 14 witchcraft > > (i) m̄-saβi 1/2 (ii) βu-saβi 14

*-cèdid- come (or go) down > -se-

*-cì 9 underneath > i-si 9 iA

*-pácà 5/6, 6, 10 twin > ma-sa 6

Its voiced counterpart is documented mostly by Class 9/10 items, which are subject to the *nj compound correspondence. However, the following evidence was elicited.

*-jàdà 5/6 rubbish-heap > [i-ɟala 5/6]

*-bààj- work wood > -βatʃ-

*-kòj- urinate > [-ɟotʃu-]

*-kòjó urine > [ma-ɟotʃo 6]

*-yɨj- come > -tʃ-

*-yɨjá DP good > -tʃa

*-yɨjud- become full > -tʃu-

We can therefore postulate the correspondence as *j > tʃ.

The nasals, as in other Bantu languages, present a clear picture with *m, *n, *ɲ > m, n, ɲ respectively. The plosive nasal compounds

are also fairly adequately documented. To illustrate:

*mp > mɓ, as in *-pángà 9/10 cave > m-ɓanga 9/10

*mb > mb, as in *-béyú 9/10 seed > m-beyu 9/10

*nt > nd̪, as in *-ntù 1/2 person > mu-nd̪u 1/2

*nd > nd, as in *-dìnd- watch over > -lindi-

*ŋk, *ŋg > ŋg, as in *-kángà 5/6, 9/10 guinea-fowl > ŋ-ganga 9/10

*-gùbò 9/10 cloth > n-guβo 9/10

*nc corresponds to s, and *nj to ʧ, both supported by the evidence of Class 9/10 nominals.

E.g. *-cádàkù$_s$ 9/10 driver or army ant > safu 9/10

*-cínj- butcher > -siʧ-

*-jókà 9/10 snake > ʧoka 9/10

2.2.3.2 Consonant reflexes in close-vowel environment

p + (-i$_s$/*-u$_s$) > f

E.g. *-pídà$_s$ 6, 14 pus > βu-fia 14

*-pù$_s$ 5/6, 7/8 stomach > ki-fu 7/8

*-píú$_s$ 7/8 knife > lu-fu 11/10a

b + (-i$_s$/*-u$_s$) > v

E.g. *-bín-$_s$ (ii) dance > -vin-

*-bú$_s$ 5 ashes > i-vu/mai-vu iA 5/6

*-bíád-$_s$ bear child > -v-

t + (-i$_s$) > s; *t + (*-u$_s$) > f

E.g. *-tíg-$_s$ leave > -siɣ-

*-tù$_s$ 14 flour > mu-fu 3/4

In the restricted environment of *t$\overset{VV}{\underset{L}{}}$ only two entries were available, namely

*-tíán- forge > -tʃan-

*-túíj- spit > -tʃw-, which does not appear to be conclusive enough evidence.

d + (-i/*-u) > r; *d + (*-ia) > ʒa

E.g. *-díbà 5, 6 milk > ma-riβa 6

*-yídìà darkness > ki-ra 7

*-dúád- wear > -rw-

*-dúud- take off (clothes) > -ru-

*-dí- eat -ʒ-

*-díá DP that, those > -ʒa

k + (-i) > s; *k + (*-u) > f; *k(*-i)+V > tʃ [1]

E.g. *-yókì 3,5 smoke > m-osi 3/4

*-kúm- come from > -fum-

*yòki-, but also *-yòki- (ii) roast > -otʃ-

g + (-i) > z; *g + (*-u) > v

E.g. *-gì 3/4 village > mu-zi 3/4

*-gùbà (i) 3 bellows > m̄-vuβa 3/4

One attesting item was elicited for the correspondence *g (*-u)+V, namely *-ŋgùà or *-yìŋgùà 3/4 'thorn' > mu-ndʒwa 3/4. This is

---

1 Correspondence also affects Class 7 independent prefix *ki+V > tʃ except in *ki+i, which follows the rule *VV > V, as in *-yídìà 'darkness' > ki-ra 7/8.

a restricted C.S., with only three entries from Saghala E.74b, Lenje M.61 and Manyika S.13a, i.e. entirely from the Eastern region.[1] On the strength of this evidence, limited though it is, we shall accept the Dawida item as valid, and a score of one point will be assigned to it.

c + (-i$_S$) > ʃ

E.g.

*-cì$_S$- grind > -ʃ-

*-cìú$_S$ 14/6 > βu-ʃu 14/6

Another piece of inconclusive evidence affects the *c + (*-u$_S$) correspondence. A sole item is available from a partial series containing only two items: *-cùg$_S$- 'wash' > (-zuɣu-) 'rub.' In absence of further supporting evidence this remains a highly tentative correspondence.[2]

j + (-i$_S$/*-u$_S$) > tʃ

*-jí$_S$ 6 water > ma-tʃi 6

*-jí$_S$- come > -tʃ-

*-jùŋg$_S$- sift > -tʃuŋg-

Although they are limited evidence, the three above items lend each other support so that the above correspondence can be postulated.

---

1 'Eastern' and 'Western' regions of the Bantu field based on topological analysis of Comparative Series. See CompB, I, 1, 61.42 - 61.45.

2 The item itself is not a valid entry of the C.S. due to its skewed meaning.

*yi > ∅, as in *-yígut- become satiated > -ɣudu-,

*-yìjá DP good > -tʃa,

*-yíjud- become full > -tʃu-,

*-yìŋgùà 3/4 thorn > mu-ndʒwa 3/4

2.2.3.3 Numerical summary

Of the total of approximately 5,000 examined word units there were 303 items with such systematic sound correspondences that they could be classified as direct reflexes of appropriate starred forms. Of these 125 were radicals and 178 were nominal stems - a fact not in disagreement with the general Bantu situation where the number of nominal stems prevails over that of verbal radicals.

Presented in a tabular fashion,

DIRECT REFLEXES OF C.B. IN DAWIDA E.74a

Verbal radicals	.....	125
Nominal stems	.....	178
Total	.....	303

Since not all correspondences could be postulated on a sufficiently large body of evidence some are more tentative than others, and some direct reflexes are more putative than others. Inferences made from statistical data should therefore be viewed with due caution.

There was also a considerable body of skewed items identified. They are dealt with in the following section.

2.2.4 Skewed reflexes of Common Bantu

In addition to the 303 direct reflexes there were in all 127 items inadmissible in the same category. They deviated from their

starred form in some respect. Deviation in shape was much more frequent than that in meaning. The prevailing tendency among the deviations in shape was their being skewed in one 'minor' feature.

It is this type of deviation especially that may be of value to diachronic interpretations in a synthesis of direct and skewed material.

Some deviations appeared to recur systematically so that a general tendency could be observed, as in *p > [h ≠].[1] Presence of another tendency helped for instance to determine whether an item was a Swahili loan. Thus lupande 'piece' is not a valid item in Common Bantu because *p > ∅ in Dawida. Since both Mvita and Unguja contain the valid item kipande 'piece' it is highly likely that we are dealing with a Swahili loan in Dawida.

There were 53 verbal radicals with a skewed shape and two with a skewed meaning, and 63 nominal stems with a skewed shape while another nine stems were skewed in meaning.

SKEWED REFLEXES OF C.B. IN DAWIDA

Verbal radicals	....	[53] (2)
Nominal stems	....	[63] (9)
Total		127

[ ] skewed in shape

( ) skewed in meaning

1 ≠ means 'does not occur in valid entries.'

2.2.4.1 Nature of individual skewings

*p

As stated above a tendency could be observed in the alternative correspondence of *p > /h/. Another correspondence which was noted several times was *p > /y/, as in *-táp- 'draw water' > /-ɗay-/.

Swahili loan words could be detected by the presence of p in corresponding Dawida words, as in *-pìm- 'measure' > /-pim-/.

Close vowel environments did not produce any skewed reflexes.

*b

The only recurring skewed reflex was *b > /ɓ/, ɓ being extraneous to Dawida.[1] It appeared in five entries of which we shall quote the following:

*-búúk- (i) rise up, wake up (ii) rise up > -βuk-, a direct reflex. But the same starred form has a skewed reflex with a specialized meaning, /-ɓuk-/ 'rise (of sun)'.

Three different skewings were noted in close vowel environments. One of them, *-bú$_{s}$ 11/10 'white hair' has a skewed reflex /imbu 5/6/ or /luimbu 11/10a/, while another C.S., although identical in shape, *-bú$_{s}$ 5 'ashes', contains a valid Dawida entry ivu 5/6.

*t

There were two kinds of deviation observed:

*t > /t/ and *t > /r/, of which t is extraneous to Dawida.

---

1 An extraneous item 'contains some feature that rules it out as a suitable entry in any part of Common Bantu.' CompB, I, 1, 32.81.

An interesting development was to be noted in C.S. 1863, which has otherwise an entirely western distribution.

*-túkù 5/6 day of 24 hours > [i-tuku 5/6]

The reflex may have arisen in the PB-A period before a complete split took place between PB-A and PB-B, and before regular sound-shifting began.[1] It is the sole departure from regular correspondences of *t in a close vowel environment.

*d

The one recurrent deviation was *d > [r] in an open vowel environment, while in a close vowel environment only [z] < *d instead of the valid r occurred more than once. It included the entry in C.S. 599 *-dúngù Cl.1/2 'whiteman' > [mzungu 1/2].

Since C.S. 599 contains only four entries all of which are from the Eastern region languages, and since Swahili is one of the three entries which involve no sound-shift, it seems likely that the Dawida item is a Swahili loan. *k > [ɣ] was the only departure noted from the regular correspondence k. The influence of a close vowel environment was documented sufficiently but quite inconclusively.

*g

Only in a close vowel environment were inadmissible entries noted. They displayed the pattern *gi > [ri], in *-gìd- 'abstain' > [-riɣid-]

*-gíkò 5/6 'fireplace' > [i-riko 5/6]

*-gínà 5/6 'name' > [i-rina 5/6]

---

1 PB-A, or Proto-Bantu A, — an ancestor language of the western section of Common Bantu; PB-B, or Proto-Bantu B, — an ancestor language of its eastern section. See CompB, I, 1, 61.81.

*c was noted as having two inadmissible reflexes, ʃ and tʃ, the direct reflex being s. No deviation was noted in close-vowel environment.

No deviations were noted among reflexes of the following:
*j, *y, *mp, *mb, *nd, *ŋk, *ŋg, *nc, *ɲ.

*nt displayed one deviation, namely ɗʃ instead of the direct reflex nɗʃ. *nj was involved in an opposite deviation. It had as its skewed reflex a nasal compound instead of the regular tʃ : *-jàdà 9 (i) hunger (ii) famine > ⌈n-dʒala 9; * → tʃala⌉.[1]

*yi
s

Two entries contained the skewed reflex i instead of the direct reflex zero. A third entry, mwiβi 1/2 'thief', displays a double skewing from its starred form *-yíbì 1/2. The valid shape would be muvi. But in view of a connected entry in C.S. 1988 *-yìb- 'steal' > -iβ-, mwiβi was not included in the above list of deviations.

---

1 * → means 'the valid shape would be.'

2.3 REFLEXES OF COMMON BANTU IN SAGHALA

2.3.1 General

The same order of items as adopted in the previous section will be used in the discussion of the second Taita language, Saghala.

First, general observations will be made on Saghala and the dialect chosen for this study. Then Saghala direct and skewed reflexes of Common Bantu will be discussed in some detail, furnished with examples from this author's data and, in conclusion, numerical evidence of the linguistic facts will be given and a summary made.

Unless otherwise stated all examples are this author's own, and are to be taken as a body of information separate from that in CompB. However, when appropriate, comparisons will be made with the information contained in CompB.

Since there are several distinct dialects of Saghala it is important to note that it is the T e r i dialect that is the object of the present investigation.[1] The data were collected in the Mlondo locality of the Teri Sub-location in 1970.

For purposes of this study 'Saghala' is used to refer to the Teri dialect.

The position of Saghala among the Test Languages as regards the number of direct reflexes of Common Bantu can be demonstrated as follows :

---

1 The others are, as identified by the Saghala speakers, Kajire, Dambi, Kishamba, Gimba, and Mgange,

1.	Bemba M.42	800 reflexes
.		.
.		.
.		.
.		.
25.	Xhosa S.41	340
26.	Tetela C.71	314
	SAGHALA E.74b	307
27.	Bulu A.74	277
28.	Duala A.24	215

Besides the 307 direct reflexes the studied material yielded an extra 121 skewed reflexes. Except for Dawida, which had a similar proportion of direct and skewed reflexes, this appears to be a much higher proportion of skewed reflexes than in the other Selected Languages. See the following comparisons.

	Dawida	Saghala	Kikuyu	Giryama	Kvita	Unguja
Direct reflexes	303	307	414	460	580	674
Skewed reflexes	127	121	63	105	166	160

As with the other languages, reflexes of nominal stems in Saghala outnumbered those of verbal radicals (190 direct, 71 skewed compared with 117 direct and 50 skewed.)

Saghala not being one of Guthrie's Test Languages, only certain selected items were included in CompB. As these do not appear to be regularly consistent with the characteristics of the Teri dialect, and in absence of data from the remaining Saghala dialects, the 182

Saghala items quoted in CompB will not be referred to in this work.[1]

2.3.2 Direct reflexes of Common Bantu in Saghala

First, sound-correspondences with C.B. will be presented in a table, which will be followed by a commentary with illustrations of individual correspondences. Where possible an example will be given from each vowel environment - open and closed, and vowel sequence, insofar as it resulted in a distinct correspondence.

Since in Saghala, as in all the other Selected Languages, the nominal stems and verbal radicals display identical sound-correspondences they will not be treated separately. Examples of data will be taken from both categories together.

Where no sound-correspondence is formulated it is to be assumed that either no evidence was obtainable or that it was inconclusive.

A correspondence postulated on evidence of less than three items will be noted in the text.

---

1 The figure of 182 reflexes is inclusive of 26 skewed items and two items from partial series.

Consonant reflexes of Common Bantu in Saghala.

(A blank in a column indicates no or inconclusive evidence.)

	+*a[1]	+*$\frac{i}{s}$	+*$\frac{u}{s}$	*-$\frac{i}{s}$+V[2]	*-$\frac{u}{s}$+V[2]	*-i+V[2]
*p	ɸ	f	f	ʃ	f	
*b	ɸ	v	v	v		
*t	ʧ	s	f	ʧ	ʧ	
*d	l	z	r	z	r	l
*k	k	s	f		f	ʧ
*g	ɣ	z	v			
*c	s	ʃ	ʃ			
*j	dʒ					
*y_1	ɸ	ɸ				
*y_2	j					
*m	m	m	m			
*n	n	n	n			
*ŋ	ŋ					
*mp	mb	f	mf			
*mb	mb		mv			
*nt	nd					
*nd	nd					
*ŋk	ŋg	s		ʃ		
*ŋg	ŋg					
*nc	s	ʃ				
*nj	ndʒ					

1 *a or any vowel other than *$\frac{i}{s}$ or *$\frac{u}{s}$.

2 Any vowel other than the preceding.

Further to the preceding table, consonants in the first and second, as well as all following positions, display the same correspondences.

The *7 > 5 vowel-shift operates in Saghala as it does in Dawida, which involves the disappearance of the distinctions *i/*i and *u/*u.[1] Thus *i, *i > i.

There is one vowel quantity in Saghala, which corresponds to two quantities in Common Bantu. It is referred to in CompB as the *VV, *V > V shift.[2]

2.3.2.1 Consonant reflexes in open-vowel environment

*p > ∅

A well attested correspondence with more than ten entries. All non-close, i.e. non-cedilla, vowels were present.

E.g. *-kúpá 5/6, 9/10 tick > ŋ-gua 9/10

*-pàká 3/4 boundary > mw-aka 3/4

*-pí- (i) become burnt (ii) become hot > -i-

*-púdá 5 wax > ma-ula 6

*b > ∅

A well documented correspondence: some thirty odd attesting items exist. However, there are some important skewed items to be noted here, and these will be discussed later.

E.g. *-bánjà 7/8, 11/10 courtyard > ki-andza 7/8

*-bíík- (i) put (ii) put away, store > -ik-

*-bón- see > -on-

*-gùbò 9/10 cloth > ŋ-guo 9/10

*-dzíbà 5,6 milk > ma-zia 6

---

1 See CompB, I, 1, 51.22

2 CompB, I, 1, 51.31.

*t > ṭ

Another well attested correspondence: some twenty items were recorded. It is interesting to note that neither Guthrie nor Wray indicate the dental articulation.[1]

E.g. *-tád- count > -ṭal-

*-té 5, 6, 11 spittle > ma-ṭe 6

*-túm- send > -ṭum-

*-gúṭà 6 oil > ma-vuṭa 6

*d > l

*d was undoubtedly the most productive starred consonant for Saghala. In our data no less than seventy items were noted as attesting this correspondence, which is far more than in any group of items attesting a Saghala-C.B. sound correspondence.

E.g. *-dàgud- practise medicine > -laɣul-

*-dìd- cry, wail > -lil-

*-dòndà 7/8 sore > ki-londa 7/8

*-pádá 9/10 kind of antelope > m-bala 9/10

*k > k

Amply attested by no less than thirty items.

E.g. *-ké 1/2 wife > mu-ke 1/2

*-kòjó urine > ma-kodʒo 6

*-kúúd- pull out > -kul-

---

1 J.A. Wray, An elementary introduction to the Taita language, S.P.C.K., London, 1894. Wray, on the other hand, does mark two different articulations of d, which do not appear complementary.

*g > ɣ

Thirty attesting items of this correspondence were recorded, e.g.

*-bògó 9/10 buffalo > m-boɣo 9/10

*-cààgud- choose > -saɣul-

*-gànà 5/6 hundred > i-ɣana 5/6

*c > s

Sufficiently attested by thirteen items, e.g.

*-cábí 14 witchcraft > u-sai 14

*-còká 5/6 axe > i-soka 5/6

*j > dʒ

E.g. *-jánì 5/6 (i) leaf > i-dʒani 5/6

*-júbà 5 sun > i-dʒua 5/6

*-kòjó urine > ma-kodʒo 6

*y_1 > ø, *y_2 > j[1]

This is the one starred symbol whose realization in Saghala is different in position C_1 and C_2. See the following two groups of examples.

*-yákà 3/4 year > mw-aka 3/4

*-yímb- sing > -imb-

*-yóg- bathe -oɣ-

---

1 I.P.A. symbol for a palatal semivowel, in Saghala examples quoted in this work represented by y.

*-béyú 9/10 seed > m-beyu 9/10

*-búy- come or go back > -uy-

*-ŋgòyì 11/10 string > lu-ŋgoyi 11/10

*-yàyò 11/10 foot > lw-ayo 11/10a[1]

*yi$_{\text{fricative}}$ > ∅

E.g. *-yígut- become satiated > -ɣuʈ-

*-yíjud- become full > -dʒul-

*-yínud- take out of water > -nul-

*m, *n, *ɲ > m, n, ɲ

These consonants provide the most unambiguous and simple correspondences with practically[2] no skewed items. We shall therefore refrain from quoting examples here.

The nasal compounds, too, display straightforward correspondences, with the exception of *nc. All their reflexes have the shape of the voiced member of each pair, e.g. *mp, *mb > mb, etc. *nj > ndʒ, but its voiceless counterpart has as its reflex s.

E.g. *-cátù 9/10 python > satu 9/10

*-yìnci 3/4 pestle > mu-si 3/4

2.3.2.2 Consonant reflexes in close-vowel environment

The relatively restricted phonetic environment of the starred close vowels affects the number of existing attesting items. A

---

1 An optional variant nyu- exists as Cl. 10 independent prefix, which is a skewed reflex of *ny.

2 The word 'practically' provides for a single exception which will be discussed further on.

'well attested' correspondence with ten entries in a non-cedilla-vowel environment is parallel to an 'extremely well attested' one in close-vowel environment. A number of correspondences had to be postulated on the evidence of two, even one attesting item, together with comparative evidence from other environments and from the structure.

p + (-i$_S$/*-u$_S$) > f

E.g. *-pígà 5/6 cooking stone > i-fiɣa 5/6

*-pìk- arrive > -fik-

*-pù 5/6, 7/8 > ki-fu 7/8

*-púkò 3/4 bag > m̄-fuko 3/4

*-pú- become fitting > -fwan-

Two items attest a correspondence in the restricted environment of *p (*-i$_S$) + V. The reflex is ʃ in both cases.[1]

b + (-i$_S$/*-u$_S$) > v

E.g. *-bí 6, 13 excreta > ma-vi 6

*-bímb- swell > -vimb-

*-bú 5 ashes > i-vu 5/6

*-búm- harvest > -vun-

Only one attesting item was elicited for the correspondence *b (*-i$_S$) + V > v, v being the reflex both of *b and *i$_S$. Thus *-bíád- 'bear child' > -val-. It is admitted into the appropriate C.S. on the supporting evidence of the preceding correspondence.

---

1 *-pɨ́éd- sweep > lu-ʃero 11/10 or 4

*-pɨ́óm- read > -ʃom-

t + (-i$_S$) > s; *t + (*-u$_S$) > f

Less than ten items document each correspondence, but their attestation is unambiguous and adequate.

E.g. *-mòtí$_S$ DP one > -mosi

*-tíg-$_S$ leave > -siy-

*-túmbí$_S$ 7/8 stool > ki-fumbi 7/8

*-túng-$_S$ tie up > -fung-

The following correspondences are highly tentative, established as they were on one attesting item each.

*-tíán-$_S$ forge > [-ʃan-] or -tʃan-

*-túíd-$_S$ spit > -tʃil-

Even though the wto items appear to lend each other support in suggesting a single reflex tʃ for both close-vowel environments, the argument concerning the first of them is somewhat involved. It concerns the shape of the Saghala item. Guthrie gives -tʃan-, Wray -ʃan- (and mʃani for 'blacksmith'), while my informants gave -tʃan- first, then corrected it for -ʃan-. The existence of a Dawida item -tʃan- is worth noting here.

We shall, however, tentatively admit tʃ as reflex of *t (*-i$_S$/*-u$_S$) +V.

d + (-i$_S$) > z; *d + (*-u$_S$) > r

Both sufficiently well attested, as in

*-díbà$_S$ 5, 6 milk > ma-zia 6

*-dìik-$_S$ bury > -zik-

*-dúud-$_S$ take off (clothes) > -ru-

*-dúg-$_S$ cook, boil > -ruy-

*-kódù$_S$ 3/4, 9/10 scar > n-goru 9/10

Identical development takes place when the starred consonant followed by a cedilla vowel is followed by another vowel. The cedilla vowel has no surface reflex in such correspondences. E.g. *-kódi- 'light fire' > -koz-.

The more frequent development is illustrated by *-dúad- 'wear' > /-rwar-/, where the first vowel does have a surface reflex. Although C_2 is skewed, the reflex in C_1 appears to be direct.

k + (-i) > s; *k + (*-u) > f

A sole item was elicited for the first correspondence but, on comparative evidence, presents no difficulties in being accepted as 'regular.' The second correspondence is amply documented and a sample only will be presented.

*-yókì 3,5 smoke > m-osi 3/4

*-cádàkù 9/10 driver or army ant > salafu 9/10

*-kúm- come from > -fum-

*-kú- die > -fw-

*-kúá 5/6 inheritance > u-fwa 14

k (-i)+ V ... undocumented; *k (*-i)+ V > tʃ

E.g. *-yádá 7/8 finger > tʃ -ala 7/8

*-yòbà 14 mushroom > tʃ -oa 7/8[1]

g + (-i) > z; *g + (*-u) > v

Two examples only were elicited to illustrate the first correspondence,

---

1 As in the other Selected Languages, Class 7 nominals in Saghala are marked by the prefix ki- (ke- in Kikuyu) < *ki-. Hence the palatalization in eg. *ki + *-yádá > tʃ ala.

the second one is somewhat better attested. No evidence is available for other close environments.

*-gi̱ 3/4 village > mu-zi 3/4

*-gi̱nà 5/6 name > i-zina 5/6

E.g. *-gu̱bà 3 bellows > mi-vua 4

*-gu̱tà 6 oil > ma-vuʒa 6

*-jògù 9/10 elephant > n-dʒovu 9/10

dʒovu 9/10

c + (-i̱/*-u̱) > ʃ

Correspondence formulated on the following evidence.

1) *-ci̱d-/*-cid- become finished > † [-sil-] [1]

-ci̱má/-ti̱má 7/8 well > † [ki-sima 7/8]

2) *-ci̱- grind > -ʃ-

*-ci̱ú 14/6 > u-ʃu 14/6

3) *-ci̱mbà 5/6, 9/10 lion > ʃimba 9/10 [2]

Because the first two entries are multivalent, i.e. they could be entered in either of the two C.S., and because the second C.S. in each case displays a valid correspondence with the Saghala entry, we shall not accept s as a direct reflex of *c in the above quoted environment.

Although we are dealing with a nasal compound in the last example, it can lend support to the evidence preceding it,

---

1 Both items in square brackets are skewed only with respect to the first starred forms quoted, which are under discussion here.

2 Cl. 9/10 have *n- as the prefix.

and, in any case, the quoted item is the only attestation of the development of that nasal compound.

No evidence was available for the correspondence of *j. The nasal consonants, as already stated in the previous section, have straightforward correspondences with no skewed items, and we shall therefore refrain from quoting examples.

Nasal compounds present the problem of restricted phonetic environment with few attesting items. In fact most lack attestation at all, and will therefore also be omitted in the following discussion.

mp +(-i$_s$) > f

*-pígò 9/10 kidney > fiyo 9/10

*-pítì 9/10 hyena > fisi 9/10

mp +(-u$_s$) > mf is attested by one entry in a partial series, *-púngàté 9 'seven' > m-fungate 9. In the light of a parallel correspondence involving the voiced nasal compound *mb, the above item has been accepted as valid.

mb + (-u$_s$) > mv is attested by a single item, namely

*-bùdà 9 rain > m-vula 9/10,

but in absence of further evidence has been accepted as valid.

Similarly, *ɲk + (*-i$_s$) > s and *ɲk(*-i$_s$)+V>ʃ have been admitted as valid correspondences. On comparative grounds the shifts appear likely but,

needless to say, are highly tentative. The following evidence is available:

*-kíngò 9/10 neck; nape > singo 9/10

*-kíá 9/10 kind of antelope > i-ʃa 9/10 iA

nc +(-í) is also attested by one item:

*-címbà 5/6, 9/10 lion > ʃimba 9/10

Together with evidence for *c + *-í this item makes it feasible for the above correspondence to be postulated.

2.3.2.3 Summary

Approximately 2,500 items were examined, of which 307 were admitted as valid entries in C.S. Of these 117 were verbal radicals and 190 nominal stems.

The single most productive starred consonant appears to be *d, which generated 79 direct reflexes in open vowel environments. All starred plosives were well documented in open vowel contexts, e.g. *k by 41 items, *g by 32, *b by 31.

Both nasals and nasal compounds lacked such abundant evidence but were adequately documented.

Close vowel environment, evidently, restricts the possibility of similarly plentiful attestation. *d followed by *í was attested by the most, namely nine, items.

To summarize:

DIRECT REFLEXES OF C.B. IN SAGHALA E.74b

Verbal radicals	.....	117
Nominal stems	.....	190
Total	.....	307

2.3.3 Skewed reflexes of Common Bantu

There was a marked predominance of skewed shapes over skewed meanings: 107 against 9. There were also five items skewed both for shape and meaning. As among the direct reflexes the skewed nominals prevailed over the verbals.

Consonantal skewing prevailed over vocalic skewing by nearly two-thirds, especially among radicals.

As noted in the preceding section, *m, *n, *ŋ represent the most straightforward correspondences in all environments.[1]

*d provided the largest number of skewed items as, indeed, it did with direct reflexes. The next largest number of skewed reflexes was that of *b, with *c, *p and *t following more or less on a par.

As for the skewed meanings, they varied widely in the degree of their departures from the connector, but such departures never appeared to be complete. Some of the substantial differences included

*kind of poisonous snake → (crocodile)

*spirit → (fool)

*bait → (point)

*soil → (clod)

2.3.3.1 Nature of individual skewings

*p

One deviation which occurred several times was *p > [p̄] instead of

---

1 Only *m yielded a skewed item, but as a multivalent entry in *-mánik- 'hang up' > † [-anik-]; & *-yánik- 'spread to dry.'

Zero. It is a distinct marker of loans from Swahili. Thus an identical starred shape could generate two reflexes in Saghala, one of them by way of Swahili.

*-pépò 9/10 wind > m-beo 9/10

*-pépò spirit > [pepo 9/10]

*b

The only recurrent deviation from the regular reflex ɸ was into β. It occurred six times, compared with 31 occurrences of the direct reflex. There were three examples of a nominal with a skewed shape in the singular and a direct reflex in plural.

E.g. *-bègà 5/6 shoulder > [i-βeɣa] / ma-eɣa 5/6

In close vowel environment one skewed item was noted, namely [-rungany-] < *-bʉ́ng- 'mix.'

*t

No recurrent deviation identified apart from two cases of *t + (*i̯) > [β] instead of the direct reflex s.

Another deviation, also in close vowel environment, involved the same item as in Dawida, namely

*-tʉ́kù 5/6 day of 24 hours > [i-ʈuku 5/6],

for discussion of which see p.48.

The rest of the skewings consisted of different 'substantial' and multiple skewings.

*d

Provided the largest number of skewed reflexes, with r occurring six

times instead of the direct reflex l and zero twice in open vowel environments.

*d ÷ *i yielded two skewed reflexes, ri and si instead of the regular zi:

*-dígì 3/4, 11/10 string > /m̄-rinsi 3/4 ; *→ m̄-zizi/

*-dígò 3/4 load > /m̄-sigo 3/4; *→ m̄-ziro/

*k

Two skewed items were identified in place of the regular s:

*-kìd- (ii) rub > /-zir-; *→ -sil-/

*-yák- build > /-ayal-; *→ -ak-/

*g

The plosive variety occurred three times instead of the fricative, which was established as the putative direct reflex. Zero occurred twice.

Followed by *i, *g had as its reflex s in one entry, namely /-siɾ-/ < *-gìd- 'abstain.' The direct reflex would be *→ -zil-.

*c yielded three different skewed reflexes, ts, tʃ and ʃ in place of the direct reflex s. Apart from ts they occurred several times each in open vowel environments, and ʃ was also noted once as a reflex of *c followed by *u.

tʃ figured in loans from Swahili, as in

*-càn- comb > /-tʃ anu-/

*-cìmb- dig > /-tʃ imb-/

Although geographically more distant, the loans are from the Zanzibar dialect rather than the Mombasa one, as the Mombasa reflex is dental ṯ.

*j and *y yielded no skewed reflexes, but *yi had twice as its reflex i instead of the regular zero.

*m, *n, *ɲ had only one skewed reflex among them, which is quoted on p.62 fn.

Among nasal compounds only *ɲk and *nc had skewed reflexes, namely

*-káá́yà 5,9a home village > ⌈kaya 9/10; * → ŋgaya⌉

*-kóndò 9/10 war > ⌈(kondo 9/10 battle; * → ŋgondo)⌉

*-cúpà 9/10 jar > ⌈ʧupa 9/10; * → sua⌉

The first of the three items is probably a loan from Giryama, where khaya is a direct reflex, while it is absent from Dawida, Kikuyu and both Swahili dialects.

2.3.3.2 Summary

Unlike in Dawida, where most skewings were 'simple' and not too different from the direct reflexes, the Saghala material yielded a rather 'unsettled' picture. There was a large proportion of 'substantial' as well as multiple skewings although the total number of skewed reflexes was less by six than in Dawida.

The following table presents the statistical results.

SKEWED REFLEXES OF C.B. IN SAGHALA

Verbal radicals	....	⌈45⌉ (3)	⌈(2)⌉
Nominal stems	....	⌈62⌉ (6)	⌈(3)⌉
Total		121	

2.4 REFLEXES OF COMMON BANTU IN GIRYAMA

2.4.1 General

The Giryama material was collected in the Mwabaya-nyundo locality of the area between Mariakani and Kaloleni in south-eastern Kenya.[1] It was later on compared with W.E. Taylor's Giryama Vocabulary and Collections, 1894, and F. Deed's Giryama-English Vocabulary, 1964.[2] The 152 Giryama entries in CompB were also consulted.[3]

In contrast to the other Selected Languages, Giryama displays remarkably little dialectal variation despite the fact that it is spoken over a considerable area about 2,500 square miles by a fairly large number of speakers-150,000. The collected data could therefore be compared with the existing works without the risk of comparing a heterogeneous body of information.

More than 5,000 items of Giryama vocabulary were examined, which yielded 460 direct reflexes of Common Bantu. Compared with the scores of the other Selected Languages, Giryama produced the third highest figure following Unguja,674, and Mvita, 580.

Comparison with the rest of the Bantu languages remains limited so far to results of Guthrie's research.[1] The table below presents the nearest higher and lower figures to Giryama's score of Common

---

1 For further details see Chapter 1, GiryamaE72a.

2 Taylor notes, almost without fail, aspiration, dental articulation, aspirated m, doubly articulated stop consonants, all of which are not noted by F. Deed.

3 Also, a short comparative wordlist was made available to me by W.H. Whiteley of some of C.B. reflexes as he worked them out.

Bantu direct reflexes. The highest and lowest scores among the Test Languages are also given.

1.	Bemba M.42	800
.		.
.		.
.		.
10.	Manyika S.13	508
11.	Rundi D.62	506
12.	Lwena K.14	485
	GIRYAMA E.72a	460
13.	Ganda E.15	450
14.	Nyoro E.11	446
15.	Nyankore E.13	441
.		.
.		.
.		.
28.	Duala A.24	215

Table: Direct reflexes of C.B.: the position of Giryama among the Test Languages.

In the following account an inventory of C.B. - Giryama sound correspondences will be presented, individual correspondences presented and furnished with examples from this authors data. Unless otherwise specified no other data will be used in the examples.

Recurring skewed sound-correspondences will also be introduced and exemplified by samples of the author's data. Finally, a numerical summary will conclude the section.

Consonant reflexes of Common Bantu in Giryama.

(A blank in a column indicates no or inconclusive evidence.)

	+ *a[1]	+ *-i (5)	+ *-u (5)	*-i+V (5)	*-i+V	*-u+V
*p	h	f	f	ʃ	ʃ	
*b	ɸ	v	v	ʒ		
*t	h	s	f	s		
*d	r/l	z	v	z	r	
*k	k	s	f	s	tʃ	kpʃ
*g	g		v			gbʃ
*c	ts	s		s		
*j	dz					
*y	ɸ[2]	ɸ				
*m	m	m	m			ŋmʃ
*n	n	ny	n	ny		
*ɲ	ɲ					
*mp	ph	f	f			
*mb	mb	v	v			
*nt	th	s	f			
*nd	nd	z				
*ɲk	kh	s				
*ɲg	ɲg	ndz	ʌ			
*nc	tsh	s				
*nj	ndz					

1 *a or any vowel other than high.

2 *y2 has as its reflex a palatal semivowel [j].

2.4.2 Direct reflexes of Common Bantu in Giryama

Individual correspondences will be presented in the following order. Firstly, sound-correspondences of consonants in open-vowel environment , secondly, sound-correspondences of consonants in close-vowel environment , namely consonants followed by *i and *u, and by a close vowel and another vowel.

As in the preceding sections, not all items attesting a correspondence will be presented. Instead one item will represent a particular environment, whenever there is an item available.

Not all sound-correspondences are sufficiently attested, and some are more tentative than others. There are also correspondences which, although theoretically likely, lack documentation altogether. These will also be reported.

There are five vowel qualities in Giryama, of which i corresponds to *i, *i and u to *u, *u, and one vowel quality. Correspondences for C_1 and C_2 are identical.

As already mentioned in 2.4.1, all examples quoted will be from this author's data unless otherwise specified. Giryama entries quoted in CompB will be noted in the text whenever they will be introduced.

2.4.2.1 Consonant reflexes in open-vowel environment

*p > h

An abundantly attested correspondence with more than thirty entries, e.g.

*-pàk- rub > -hak-

*-pàká 3/4 boundary > mu-haka 3/4

*-píni 3/4 handle > mu-hini 3/4

*-pód- become cured, get well > -hol-

*-pùdud- strip off leaves or skin > -hulul-

*b > ∅

Another well attested correspondence, with some twenty entries, e.g.

*-bádik- become split > -arik-

*-bícì DP unripe; uncooked > -itsi

*-bíík- (i) put (ii) put away, store > -ik- (i)

*-bòd- become rotten > -ol-

*-búúdi- ask > -uz-

*t > h

Another correspondence with h in Giryama, and equally well attested as the first one. There were 15 occurrences noted.

E.g. *-tádà 5/6, 7/8, 11.10 platform > mu-hala 3/4

*-té 5, 6, 11 spittle > ma-he 6

*-tí 7/8 stool > ki-hi 7/8

*-túm- send > -hum-

*d > r/l

*d > l except when preceded or followed by a front vowel, in which case it became r.

Both correspondences are well attested by at least twenty items each. E.g.

*-dàdá 3/4 kind of palm tree > mu-lala 3/4

*-dàdú madness > vi-lalu 8

*-dóbò 5/6, 7/8, 9/10 fish-hook > ki-loo 7/8

*-kód- choke > -kol-

*-dáádik- invite > -larik-

*-dè DP long > -re

*-dèèdó today > rero

*-dìmà 7/8 hill > mu-rima 3/4

*k > k; *k (*-u)+V > kp[1]

The first is a simple correspondence attested by some fifty items. k appears to have been the most productive starred consonant for Giryama. As the direct reflex is identical with its starred symbol, no examples will be given of the correspondence.

The second correspondence, although by no means as abundantly attested as the first, is sufficiently documented by four items. It is characteristic of other 'Mijikenda languages' also.

E.g. *-kúápà 5/6, 9/10 armpit > kpaha 5/6

*-kúéd- go up > -kper-

*g > g; *g +(*e/*i) > dʒ; *g (*-u)+V > gb[1]

First two correspondences are sufficiently documented, but the third required further comparative evidence to support it (see *k (*-u)+V, *m (*-u)+V).

E.g. *-gànà 5/6 hundred > gana 5/6

*-gònò 3/4 fish-trap > mu-gono 3/4

*-gèdi- try > -dʒez-

*-gèni 1/2 stranger > mu-dʒeni 1/2

*-gí 5/6 egg > i-dʒi 5/6

*-gù- fall > -gb-

*-góyè 11/10 string > /lu-gbe 11/10; ← *guye/

In spite of being skewed in one vowel the last item demonstrates the valid correspondence *g (*-u) +V > gb.[2]

---

1 kp, gb, ŋm are doubly articulated stop consonants (or velarlabials) which correspond to the Swahili kw, gw, mw.

2 The correspondence operated probably in two stages, with *y disappearing in the first stage.

*c > ts

Well attested by, among others, the following items.

*-cábí 14 witchcraft > u-tsai 14

*-cèk- laugh > -tsek-

*-cìmb- dig > -tsimb-

*-còká 5/6 axe > tsoka 5/6

*-cúngù 14 poison > u-tsungu 14

*j > dz

Well attested, i.e. by more than ten items. E.g.

*-jàdà 5/6 rubbish-heap > dzala 5/6

*-jì 3/4 village > mu-dzi 3/4

*-júbà 5 sun > dzua 5/6

*y_1 > ∅; y_2 > j[1]

The former attested by ten, the latter by eight items.

E.g. *-yákà 3/4 year > mw-aka 3/4

*-yéd- shine > -er- spec

*-yíg- imitate > -ig-

*-yóg- bathe > -og-

*-yàyò 11/10 foot > lw-ayo 11/10

*-yòyò 3/4 heart > m-oyo 3/4

*-búy- come (or go) back > -uy-

*-béyú 9/10 seed > m-beyu 9/10

*yi > ∅

Although the above starred sequence has a number of putative skewed reflexes the above correspondence was possible to establish. The

---

1 I.P.A. symbol for a palatal semivowel, in Giryama examples in this work represented by y.

more important skewed reflexes will be discussed further on. The zero reflex was attested by

*-yícò 5/6 eye > dzi-tso 5/6

*-yícòdì 3/4, 4, 5/6, 6 tear from eye >†tsozi[1]

*-yìj- come >†-dz-[2]

*-yìkad- (i) sit (ii) dwell; be >†-kal-[3]

*-yínà 5/6 name > dzi-na 5/6

*-yínò 5/6 tooth > dzi-no 5/6

--yínò 5/6 tooth > dzi-no 5/6

--yìnci 3/4 pestle > [mu-tsi 3/4; * >mu-tshi]

n--yíngid- come (or go) in > -ngir-, [-angir-]

The starred nasal consonants, as would be expected according to evidence from other Bantu languages, provided the simplest correspondences both in open and close vowel environment. We shall therefore refrain from exemplifying the simple correspondences.

The last correspondence is supported by a single item together with comparative evidence (see similar development in *k (*-ʊ)+V and *g (*-u)+V),

*-mùàg- scatter, sprinkle > -ɲmag-

Nasal compounds displayed straightforward correspondences. The pattern was as follows:

Those with second component voiceless corresponded to that second compnnent aspirated. Thus *mp > ph, *nt > th, *nk > kh, *nc > tsh.

---

1 multivalent with *-códì.

2 possibly multivalent with *-jí-.

3 multivalent with *-kàd-.

Those with second component voiced corresponded simply to the compound unchanged. Thus *mb > mb, *nd > nd, *ng > ng, *nj > ndz.[1]

E.g. *-pácà 5/6, 6, 10 twin > phatsa 10

*-ntù 1/2 person > mu-thu 1/2

*-kánjgà 5/6, 9/10 guinea-fowl > khanga 9/10

*-cá (9/10) point > tsha 9/10

*-bónò (9/10) castor-oil bean > m-bono 9/10

*-ndà or *-yìndà 3/4 garden > mu-nda 3/4

*-gànò 7/8, 9/10, 11/10 tale > n-gano 9/10

*-jàdà 9 (i) hunger (ii) famine > n-dzala 9/10

2.4.2.2 Consonant reflexes in close-vowel environment

The relatively restricted environment yielded, as with the other Selected Languages, substantially less evidence than the much wider open-vowel environment.

Among the starred consonants the stops were best attested with at least three items in each environment, i.e. followed by high front and high back vowel. *k and *d were abundantly attested by five to fifteen items for each relevant environment.

But *c and *j especially, as well as nasal compounds, often lacked attestation however theoretically possible the putative direct reflex seemed to be.

Where there are three or less items available the evidence will be given in full, which will be indicated by absence of the abbreviation 'e.g.' preceding the examples.

---

1 Since *j represents Giryama dz it can be included in the statement.

p + (-i̯/*-u̯) > f; *p /(*-i̯) + V, (*-i) + V/ > ʃ

E.g. *-pí̯gà 5/6 cooking-stone > figa 5/6

*-pú̯kò 3/4 bag > mu-fuko 3/4

*-pí̯éd- sweep > -ʃer-

*-pí̯ó́m- read > -ʃom-

*-pí́- become burnt > -ʃ-

The last example is the only evidence for the correspondence.

b + (-i̯/*-u̯) > v; *b (*-i̯) + V > ʒ

*-bí̯ 6,13 excreta > ma-vi 6

*-bú́n- harvest > -vun-

*-bú́nj- break, snap > -vundz-

The first correspondence is further attested by the Class 8 independent prefix in consonantal stems, namely *bi- > vi-.

The reflex ʒ was in our data attested by a single item -ʒal- < *-bí̯á́d- 'bear child', and also by Class 8 independent prefix in vocalic stems, as in

*-bú́nù 7/8 waist > tʃ-unu/ʒ-unu 7/8

The sound-correspondence affecting the last item appears to operate in two stages. In the first stage *b has as its regular reflex zero. In the second stage the Class prefix *bi̯ followed by the back vowel gives rise to the voiced palato-alveolar fricative. In other words,

*bi̯ - bú́nù > *bi̯-unu > ʒunu

t + (-i̯) > s; *t +(*-u̯) > f

E.g. *-tì̯kí́ 7/8 stump of tree > ki-siki 7/8

*-tɡ́án- forge > -sany- [Y]

*-tɡ́úd- forge > -ful- spec (work iron)

*-tɡ́úmɔ̀ 5/6 spear > fumo 5/6

The second item attests identical correspondence for *CV and for *CVV sequence.

d + (-i) > z; *d + (*-u) > v

E.g. *-dɡ́ 11 string > lu-zi 11/10

*-dɡ̀ɛ̀m- sink > -zam-

*-dɡ̀ùd- pull up > -zul-

*-dɡ́úŋgù 1/2 whiteman > mu-zungu 1/2

*-dɡ́úd- take off (clothes) > -vul-

*-dɡ̀ugud- open > -vugul-

*-dɡ̀ut- pull > -vuh-

*-dɡ́ɑ́d- wear > -val-[1]

k + (-i) > s; *k + (*-u) > f; *k (*-i) + V > tʃ

E.g. *-kɡ́dɡ̀ soot > mi-sizi 4

*-kɡ́ndɔ̀ 3/4 noise > mu-sindo 3/4 spec

---

1 This author's data are at variance both with W.E. Taylor, F. Deed and W.H. Whiteley (Guthrie did not include the item in CompB.) All three quote -vwal- 'wear' and -vwik- 'clothe,' the latter corresponding to *-dɡ́ík-. The present author considers her own data as likely to be contaminated by Swahili because the Swahili reflexes of *-dɡ́ɑ́d-, *-dɡ́ík- are -va-, -vik- respectively for both Mvita and Unguja. -viʃ- was also elicited as the causative form, with extra palatalization. However, there is a comparable correspondence involving *g, namely *-gúí/ *-gúí 3/4 'arrow' > mu-vi 3/4, which would appear to lend some support to this author's entries quoted above as being valid.

*-kṵ́- die > -fw-

*-kṵ́tà 6 oil > ma-fuha 6

-fw- is the only reflex of *k followed by a high back vowel and another vowel in which both vowels have a surface reflex. A parallel correspondence has been noted for *d as attested by Taylor's and Whiteley's evidence. (See previous page.)

The third correspondence, where V ≠ *i or *i̯, was attested by the following items among others:

*-yòki- burn > -otʃ-

*-kí- dawn > -tʃ-

*-yádá 7/8 finger > tʃ-ala 7/8

*-yòmbò 7/8 boat > tʃ-ombo 7/8

*-yùdá 7/8 kind of frog > tʃ-ula 7/8

Note that the last three items demonstrate the process affecting Class 7 independent nominal prefix, stem vocalic.[1]

g + (-i̯) > z; *g + (*-u̯) > v

*-gi̯d- abstain > -zir-

*-gi̯mà DP whole > -zima

*-gu̯bò or *-bu̯bò or *-dʒu̯bò 4 bellows > mu-vuo 3/4

*-gu̯nd- ripen artificially > -vundy- [ɣ]

*-gúí/ *-gúí 3/4 arrow > mu-vi 3/4

The last item is the only exemplification available for the correspondence of *g followed by a high back vowel and another vowel. It is a highly putative correspondence, especially in view of the

---

1 As in, e.g., *ki-+*-yádá > ki- +-ala > tʃ-ala.
Note that *y represents zero in Guthrie's notation.

parallel correspondence affecting *d which is likely to be not valid. (See correspondences of *d in close-vowel environment.)

c + (-i) > s; c + (*-u) not attested

Correspondence attested by two items for the *-ciC- environment and by three items for the *-ciV- environment. Both have the same reflex.

E.g. *-címa 7/8 well > ki-sima 7/8[1]

*-ciag- or *-tiag- grind > -sag-[2]

*-ciád- remain > -sal-

*-cìó 14/6 face > u-so 11/10

j + (-i) not attested; *j + (*-u) > ts ?

Only one entry was elicited for *ju. The question mark following the putative reflex indicates that further evidence may well produce a different reflex. The existing correspondence is *-jùng- 'sift' > -tsung-.

The overall pattern of nasal compounds in close-vowel environment, more or less adequately attested, reveals correspondences identical with those affecting the non-nasal part of each compound. Thus *p + (*-i/*-u) > f, *mp + (*-i/*-u) > f, or more concisely, *p/*mp + (*-i/*-u) > f; *b/*mb + (*-i/*-u) > v, etc.

As in other Bantu languages, in C_1 position in stems, the development of the nasal compounds affects by and large nominals of Class 9 and 10 only as the nasal is the marker of the two Classes.

---

1 Multivalent with C.S. 1753 *-tíma.

2 The two starred forms are contained in the same partial series No.99. It is not tone-marked.

*ndu, *nku, *ncu, *nji/*nju lack evidence altogether, however theoretically possible their reflexes appear to be, while some correspondences are attested by less than three items. In the latter case the evidence will be quoted in full.

mp + (-i/*-u) > f

*-pítì 9/10 hyena > fisi 9/10

*-púngàté 9 seven > fungahe 9

mb + (-i/*-u) > v

*-bìtá 5/6, 9/10 war > viha 9/10

*-búdà 9/10 rain > vula 9/10[1]

nt + (-i) > s; *nt + (*-u) > f

*-tíkù 14/6 night > u-siku/siku 11/10

*-tíngà 11/10 bow-string > u-singa/singa 11/10

*-tùngò 9/10 civet-cat > fungo 9/10

Only on structural evidence could the following correspondence be posited, however tentatively.

nd + (-i) > z,

as in *-dìtò DP heavy > -ziho. Dependent prefix in Classes 9 and 10 in Giryama is zero.

nk + (-i) > s

As the preceding nasal compounds in close-vowel environment, *nk has a correspondence identical with that of its second component. However, only one attesting item was identified among our data, namely

---

1 See fn.1 on p.82 .

sinɟgo 9/10 < *-kíɲgò 9/10 neck.

ɲg +(-i) > ndz; *ɲg + (*-u) > v

The first correspondence was posited on the evidence of two items, the second on the evidence of a single item. By analogy with its voiceless counterpart as well as other nasal compounds it would appear that the *ɲgi>ndzi correspondence would require further evidence for it to be unambiguous.

(By analogy the reflex would have been nzi.)

*-gì 9/10, 11/10 fly > in-dzi iA

*-gìgè 9/10 locust > n-dzije 9/10

*-gùbú 9/6, 9/10 hippopotamus > vuu 9/10[1]

A single item attests the correspondence *nci > si, namely simba 9/10 > *-címbà 5/6, 9/10 'lion', while a correspondence involving a high back vowel can only be assumed to be theoretically likely.

No evidence was recorded for *nji/*nju.

*yi > ∅

*-yícò 5/6 eye > dzi-tso 5/6

---

1 v, also encountered in vula 9/10 < *-búdà 9/10 'rain,' is not dissimilar in its articulation from Guthrie's Labio-dental (2) (CompB, II, 3, p.16: ... 'the inside of the protruded lower lip is involved rather than the edge of the lip, and there is also lip-rounding.')
The Ciryama v has also considerably less friction than the 'plain' labiodental v.
More evidence would be needed in order to establish unequivocally whether v is the regular reflex of the sequences N+voiced plosive when followed by a high back vowel.

-yícòdì/-còdì 3/4, 4, 5/6, 6 tear from eye > tsozi 5/6

*-yìj- come > -dz-

*-yínà 5/6 name > dzi-na 5/6

*-yínò 5/6 tooth > dzi-no 5/6

*-yìnci 3/4 pestle > [mu-tsi 3/4; *-> mu-sthi]

*-yíngid- come (or go) in > [-angir-], -ngir-

Apart from the above evidence there are several important skewed reflexes, which will be discussed further on.

2.4.3 Skewed reflexes of Common Bantu in Giryama

The overall figure arrived at was 105 skewed items. Of these 16 were skewed for meaning, 85 for shape, and four for both meaning and shape. This, in comparison with the other Selected Languages, is not an unexpected result. There were 460 direct reflexes established in Giryama, and it appears 'normal' for a Selected Language to have its indirect reflexes numbering between approximately one fifth and one third of that number.

However, Giryama did have proportionately less skewed items than Dawida and Saghala. Compare figures in the following table.

	Dawida	Saghala	Giryama
Direct reflexes	303	307	460
Skewed reflexes	127	121	105

As in Dawida, and unlike in Saghala, most skewed items involved 'simple' skewings with one only feature skewed. As expected, consonantal skewings prevailed over vocalic ones, and there were

certain pronounced tendencies noted among the skewings.

Provided that a tendency could be discerned among the skewings it will be introduced and illustrated by examples. Where three or less items support a suggested tendency they will be quoted in full.

2.4.3.1 Nature of individual skewings

*p

Out of twelve skewed items there were seven with a *p > [p] correspondence. On comparative evidence all the items concerned may be said to be Swahili loans. Out of the seven C.S. involved only one contains an entry from the other Selected Languages apart from Giryama and Swahili. In two of the C.S. the evidence is clearer still. The nominal and verbal denoting an unambiguously 'cultural term', 'measure', are members of C.S. with little or no phonological variation. There are[a] further four skewed items appended to the C.S. which display equal homogeneity. For exemplifcation the two C.S. are given in full in a footnote.[1]

---

1 C.S. 1519 *-pìm- measure

G.23	Sambaa	-him-	N.31	Nyanja	-pim-
G.42d	Unguja	-pim-	P.21	Yao	-pim-
L.33	Luba-Katanga	-ɸim-	S.13a	Manyika	-pim-
M.42	Bemba	-pim-			

[E.72a	Giryama	-pim-;	p ≠]
[F.21	Sukuma	-phiim-;	ph≠]
[S.21	Venda	-p'im;	p'≠]
[S.51	Tswa	-pim-;	p ≠]

C.S. 1520 -pìmò 7/8 measure

		Cl. 7/8
G.42d	Unguja	ki-pimo
G.44a	Ngazija	ʃi-βimo
L.33	Luba-Katanga	ki-pimo
M.42	Bemba	itʃi-pimo
P.21	Yao	ci-pimo

1519: These entries could be attributed to a source-item in PB-B, but as the skewed items in the footnotes seem to be loan-words, probably from Swahili G.42, some of the apparently valid entries might in fact also be intrusions.
CompB, II, 4, p.57.

No further tendencies could be discerned among the rest of the skewed items. No skewing was encountered of the correspondences *p + (*-i/*-u) > f, *mp > ph.

*b

The sixteen skewed entries had two patterns repeated among them, namely *b > ⌈b⌉ and *b > ⌈h⌉. The first recurred three, the second four times. The *b > ⌈b⌉ correspondence involves items all of which have corresponding entries in Swahili (all three in Unguja, two only in Mvita), thus indicating they are loans from Swahili. b is extraneous to Giryama.

Thus, for example, *-bánd- 'split' tr ⌈-bandzul-⌉ in Giryama and ⌈bandu ph p⌉ in Unguja and Mvita.

The four instances of the skewed reflex h appear to be more complex, especially in connection with the irregular reflex of $*y_1$ (the regular being zero), which turned out to be h in one entry. The following are the four Giryama skewed items.

*-báb- itch > ⌈-hah-; h → *p/*t⌉

*-bád- marry > ⌈-hal-; h → *p/*t⌉

-bìdí/-bídì DP two > ⌈-hiri; h → *p/*t⌉

*-bìnd- plait > ⌈-hind- coil, wind; h → *p/*t⌉

The item involving the skewed correspondence $*y_1 > ⌈h⌉$ is this:

*-yàní la/2 who? > ⌈hani; $*y_1 > \phi$⌉

On the available evidence it would appear that the former four items suggest a development *b > ϕ > h provided that, as in the present case, the likelihood of contamination from Swahili is excluded: correspondences between Swahili and Giryama do not involve y → h.[1]

---

1 The only case of *b > h correspondence encountered in CompB was recorded in the following note of Tiv Z.1: 'Several instances of certain skewed sound-shifts occur, such as ⌈...⌉ *b > h ⌈...⌉. CompB, I, 2, p.64.

The *mb > mb as well as the close-vowel environment correspondences yielded two and five skewed items respectively but there was no tendency apparent.

*t

Two items were identified as skewed for meaning, eleven skewed for shape. Eight of them contained the irregular ʇ as a reflex of *t. In contrast to these skewed dental ʇ's there were 33 direct reflexes, namely h. The following are the eight irregular entries with ʇ.

Note: Following each Giryama entry a concise statement is given concerning existing entries from other Selected Languages. E.g. ' + Sw.' means that reflexes of the relevant item exist also in both Swahili dialects; no information indicates that a reflex was elicited in Giryama only.

*-tàbì 3/4, 5/6, 7/8 branch > [ki-ʇai 7/8, lu-ʇai 11/10] + Sw.

*-tád- count > [-ʇal-] + all but Sw.

*-tàdò number > [ʇalo 5/6] + D

*-táp- pillage > [-ʇav-]

*-tém- cut > [-ʇem-] + all

*-téndé kind of palm-tree > [mu-ʇende 3/4] + Sw.

*-tóód- pick up > [-ʇol-]

*-yáto 14/6 canoe > [(mw-aʇo 3/4 beehive)]

For possible conclusions regarding the status of [ʇ] each of the above C.S. would merit a special analysis with reference to the geographical spread of Guthrie's C.S. Entries from other Selected Languages do not lend themselves to any hypotheses.

While the *-CVC- pattern involving $*t_1$ did not yield any skewed reflexes, the -CVVC- and -CVVC- patterns presented rather a complex picture:

*-tíúd- or *-kíúd- rub > [-tsul-]

*-túíj- spit > [-tsuh-]

*-tùùd- swell > [-vw-]

*d

The most frequently recurring skewing was *d > [r] in an environment where l would be expected. (*d > r when preceded or followed by i or e.) From a total of 15 items skewed for shape seven were subject to this correspondence. E.g.

*-dàndú 3/4, 5/6 debt > [randu 5/6]

*-dòŋgò 3/4, 4 ten > [mu-roŋgo 3/4]

*-dùŋg- become right; become straight > [-ruŋg-]

There may be several explanations of this predominance of r, of which contamination from adjacent languages could be one.[1] Different dialectal developments could be another reason if synchronic evidence were available of dialectal variation. Influence of Cl.5 prefix (always a front vowel) might account for the correspondence in the first item if there were not the balancing evidence of other Cl.5 items, which attest *d > l.

No tendencies were discerned among the skewed reflexes of *d in close-vowel environment or when preceded by a homorganic nasal.

---

1 D. Nurse, G. Philippson, The North-Eastern Bantu Languages of Tanzania and Kenya: a classification, The Institute of Swahili Research, University of Dar es Salaam, 1974. Digo E.73 has regularly *d > r, among Digo E.73, Chonyi E.72c, Pokomo E.71. No correspondences given in CompB for Digo and Chonyi.

*k

The evidence for what we decided to establish as a putative skewed reflex, is by no means unambiguous. The direct reflex is k in the context *ka, *ko, *ku. Followed by a front vowel, *k corresponded to tʃ ten times, to k seven times.

*ke > 7 tʃe, 1 ke

*ki > 3 tʃi, 6 ki

An additional piece of evidence is provided by the reflex of Cl.7 prefix *ki > ki. (Followed by a front vowel and another vowel, *k has a regular reflex tʃ.) Against the evidence of Cl.7 prefix there is the evidence of a parallel development of *g + (*-e/*-i) into an affricate, but the former evidence appears to be somewhat stronger. Whereas parallel developments are not to be found regularly, the C.S. of concord prefixes display great regularity.[1]

None of the remaining Selected Languages yielded a similar correspondence, skewed or direct, except for a single item in Unguja, /(ma-tʃeo 6 dawn)/ < *-kéédó 11 'morning'.

We have tentatively treated tʃ as skewed in the context *ke/*ki.[2]

*g

The heterogeneous character of the skewings would make it unjustified to attempt a deduction of any tendencies among them. Only the palatal semivowel and zero occurred twice each.

E.g. *-tíg- leave > /-siy-; * → -sig-/

*-gòngò 3/4 (i) back (ii) backbone > /m-ongo/mi-ongo 3/4; * → mu-gongo/

---

1 CompB, Comparative Series of concord prefixes, II, 4, 220-228.

2 Nurse and Philippson, op.cit., claim the existence of regular palatalisation of *k to tʃ before *i, *e, both for Giryama E.72a, Chonyi E.72c, Digo E.73, and even Standard Swahili. See the section on Unguja further on.
Guthrie gives only one reflex of *k in Giryama, namely k.

As with *k, no skewed reflexes occurred involving close-vowel environment of *kuV, *guV, or the nasal compounds.

*c

No clear pattern could be established among the four skewed items. Two of them, however, suggest Swahili contamination since both entries contain tʃ , the regular reflex of *c in Unguja. C.S.291 item is an apparent loan word: as a cultural term, 'bead', and a homonym in all four entries, uʃaŋga, the evidence is overwhelming.

In other environments no skewed reflexes occurred.

*j

Three skewed reflexes were from osculant C.S.

*-jègò 5/6 tooth > [† dʒego 5/6; † *-gègò]

*-jèmbè 5/6 hoe > [† dʒembe 5/6; † *-gèmbè]

*-yìj- come > [† -dʒ-; † *-jí-]

Other environments yielded no skewed reflexes.

*y

Another incoherent picture here: the three identified skewings involve one skewed vowel, one correspondence $*y_1 > [h]$,[1] and $*y_1 > [dʒ]$ in a C.S. osculant with *-gèdi-.

The three nasal sounds *m, *n, *ɲ yielded no skewed reflexes.

---

1 C.S.1925 *-yàní 1a/2 who? > [hani; * → ani]

2.4.3.2 Numerical summary

SKEWED REFLEXES OF C.B. IN GIRYAMA

Meaning skewed	16
Shape skewed	85
Both skewed	4
Total	105

2.5 REFLEXES OF COMMON BANTU IN KIKUYU[1]

2.5.1 General

In the following section a statement will be made concerning the relationship of Kikuyu and C.B. as it is documented in CompB. Kikuyu, being one of Guthrie's Test Languages, has been subject to a detailed examination lexically and morpho-phonemically. As far as we know,[2] Guthrie's data comprised his own field notes, A.R. Barlow's Tentative Studies in Kikuyu Grammar and Idiom, Edinburgh 1914, and L.E. Armstrong's The Phonetic and Tonal Structure of Kikuyu, C.U.P., 1940. The locale of Guthrie's research is not stated in CompB or elsewhere.

Below we shall elaborate on the sound-shift inventory as worked out by Guthrie,[3] on direct reflexes of the Common Bantu reconstructions and, in a separate subsection, observations will be made on the character and numbers of Kikuyu reflexes that do not display regular correspondences with their Common Bantu reconstructions or starred forms.

All examples are taken from CompB unless otherwise stated, but all numerical statements and computations are this author's responsibility.

---

1 In this work the traditional term 'Kikuyu' has been used for the correct Gikuyu or, as in CompB, Gekoyo which indicates the existence of seven vowel qualities in the language. By Dahl's law of dissimilation, of two voiceless consonants in two successive syllables the first becomes voiced. In Kikuyu this rule operates only with $*k_1$.
When referring to the E.50 group of languages in CompB, Guthrie uses the form 'Kikuyu-Kamba.'

2 Personal communication.

3 CompB, I, 2, p.46.

With its 414[1] direct reflexes of Common Bantu Kikuyu is in the seventeenth place when compared with the other twenty seven Test Languages. The nearest higher figures belong to Mbundu R.11, 427 reflexes, Nyankore E.13, 441, and Nyoro E.11, 446, the nearest lower belong to Kamba E.55, 402, and Sotho S.33, 402, followed by Bobangi C.32, 401. Below, these figures are presented tabularly. For the complete list see p.33 .

1.	Bemba M.42	800
	.	.
	.	.
	.	.
14.	Nyoro E.11	446
15.	Nyankore E.13	441
16.	Mbundu R.11	427
	KIKUYU E.51	414[1]
17.	Kamba E.55	402
18.	Sotho S.33	402
19.	Bobangi C.32	401
	.	.
	.	.
	.	.
28.	Duala A.24	215

Table: Direct reflexes of C.B.: the position of Kikuyu among the Test Languages.

---

1 M. Mann's figure is 424, and was based on Guthrie's Register of Entries from the Test Languages, CompB, I, 2, 65-104. (From personal communication from M. Mann.) This author's figure was based on the count of the actual entries in Common Bantu.

Of the 414 direct reflexes 191 were radical reflexes while 223 were reflexes of nominal stems. As with Dawida, Saghala and Giryama, this is a representative situation of the Bantu field.

The types of sound-correspondences found in the verbal radicals and nominal stems are identical, and remarkably straightforward in contrast to the other Selected Languages.

This fact is to some extent due to the existence of seven vowel qualities in Kikuyu which display one-to-one correspondence with the seven vowel qualities of Common Bantu.[1] Kikuyu has two vowel quantities; their correspondences with Common Bantu starred forms are governed by the structural rules of Kikuyu. It is the realizations of C.B.-Kikuyu correspondences which are presented as data in CompB in the 'narrow' phonetic transcription.

As in CompB, the actual entries in the C.S. are not tone-marked, but it is worth noting that Kikuyu has a 'clear' tone system,[2] which means that its tones display one-to-one correspondence with the tone marks of appropriate starred forms. In Kikuyu 'the distinctive tone of each syllable occurs on the following syllable' which, however, 'does not affect the tonal correspondences which are quite regular and complete in this language.'[3]

---

1 This is not the only existing state of affairs with languages that have a seven vowel qualities system. E.g. Nz ɛ bi B.52 is a seven-vowel language whose vocalic system is not directly correlated with C.B. Thus e.g. *e > ɛ /(*-i, *-i) > e/; *i > i; *i > i.

2 Guthrie's term. For detailed discussion of Guthrie's tonal data see H. Carter, 'Tonal data in COMPARATIVE BANTU,' ALS, XIV, 1937, 36-52.

3 CompB, I, 1, 51.51.

E.g. *-búdag- kill > -ọrá ɣ-

*-búdị̀ 9/10 goat > m-bọ̀rí 9/10 [1]

2.5.2 Direct reflexes of Common Bantu in Kikuyu

The table on the following page is an elaboration of the sound-correspondence inventory for Kikuyu presented in CompB. For the sake of comparison we shall reproduce it here.

*p/*mp > h/.h; [2] *b > ∅ /(*-ị̀a) > t͡ʃ; *t/*d > t/r;

*nt,*nd > nd; *k > k/(*~ -ọ) > ɣ; *g > ɣ; *ŋk, *ŋg > ŋg;

*c/*nc > ɟ/.ɟ; *j/*nj > ∅ /nj

There are also Cl.7 and 8 independent prefixes quoted, but these are direct reflexes of the starred forms given above.

Vowel correspondences can be represented as follows:

*7V > 7V, meaning that *a > a, *e > ɛ, *i > ẹ, *ị̀ > ị, *o > ɔ, *u > ọ, *ụ > ụ in the orthography of CompB. [3]

Kikuyu not being affected by any complex sets of rules governing its sound-correspondences with Common Bantu it is possible to discuss in the following section correspondences both in open-and close-vowel environments. Unlike in most five-vowel languages the influence of *ị̀ and *ụ does not bring into motion any mechanism of intricate sound developments.

---

1 Tone-marking of Kikuyu items carried out according to T.G. Benson (ed.), Kikuyu-English Dictionary, C.U.P, 1964. 'Guide to the pronunciation indicators following the entries,' xiii-xlix, prepared by A.E. Sharp, contains tables with exemplification of tone groups.

2 A dot before an initial consonant indicates potential vowel length: thus the short final vowel of a preceding word is lengthened. (Guthrie's convention.)

3 T.G. Benson's Kikuyu-English dictionary, 1964, follows a different convention.

Consonant reflexes of Common Bantu in Kikuyu.

$C_1 = C_2$ [1]

	+ V; V = any vowel
*p	h
*b	∅; *bi̯a > tʃi̯a
*t	t
*d	r
*k	k
*g	ɣ
*c	ɔ
*j	∅
*y	∅; *yi̯ > i̯
*m	m
*n	n
*ɲ	ɲ
*mp	.h [2]
*mb	mb
*nt	nd
*nd	nd
*nk	ng
*ng	ng
*nc	.ɔ
*nj	nɟ

1 Sound-correspondences are identical for the first and second consonantal positions.

2 A dot before an initial consonant indicates potential vowel length: the short final vowel of a preceding word is thus lengthened. (Guthrie's convention.)

Most of the postulated sound-correspondences are attested by a sufficient number of items, which in our context means at least four. A 'well documented' sound-correspondence, as stated in an earlier section, is one with at least ten attesting items.

Contrary to experience with the other Selected Languages it was not unusual in Kikuyu to reach figures above ten when eliciting attesting items for a sound-correspondence, because the division into close- and open-vowel environments, producing differing correspondences, did not prove necessary.

In the following text not all available attesting items will be presented. Examples will be given from the data which are considered to be as most clearly illustrative of a particular sound-correspondence.

*p > h

Attested abundantly and unambiguously by 49 items, e.g.

*-pàk- rub >-hak-

*-pèni 9 lightning >rọ-hɛni 11

*-pic- hide >-hiɔ-

*-pùkò 3/5 bag >mọ-hukɔ 3/4

*b > ϕ

Altogether thirty odd items attested this correspondence. Owing to a substantial number of nominals in Cl.9/10, whose prefixes are marked by a homorganic nasal, and which are subject to a separate correspondence, there were fewer entries under the *b correspondence than under the *p > h one with fewer Cl.9/10 nominals.

E.g. *-bàki- light fire > -aki-

*-bóp- bind > -ɔh-

*-bín- dance; sing > -in-

*-bún- break, snap tr > -un-

There is one *b correspondence that has a voiceless affricate as its reflex, but only one attesting item was encountered.[1]

*b (*ia) > tʃ ,

as in *-biád- bear (child) -tʃiar- .

However, the correspondence is further attested by identical development in Class 8 independent prefix *bi- when followed by a vowel stem:

*-yánà 1/2 child > mw-ana/tʃi-ana 1/8

*-yùdá 7/8 kind of frog > ke-ora/tʃi-ora 7/8

*-jádá 7/8 finger > ke-ara/tʃi-ara 7/8

*t > t

Forty odd items attest this correspondence, e.g.

*-tú 5/6, 15/6 ear > rɔ-tɔ 15/6

*-tíg- leave > -tiir-

*-tù (5/6) cloud > i-tu 5/6

*-dóót- dream > -rɔɔt-

*d > r

Over sixty items attest this most straightforward correspondence. *d presents a rather complex picture in the other Selected Languages and, from this point of view, separates Kikuyu from the remaining languages of this work.

---

1 This is contrary to Guthrie's criterion of a 'regular' correspondence, which should be attested by at least three items. Cf. CompB, I, 1, 22.31.

E.g. *-dámb- lick > -ramb-

*-dègid- become slack > -rɛɣɛr-[1]

*-díá DP that, those > -rẹa

*-dʒùk- revive > -rɪ̀ọk-

*-dʒúg- cook, boil > -rʊɣ-

*k > k

Again, abundantly attested by, among others,

*-kám- milk > -kam-

*-kìd- grind > -kir-

*-kúd- grow up > -kọr-

*-kúndò 5/6, 7/8 knot > i-kūndɔ 5/6

Due to the operation of 'Dahl's law'[2] in Kikuyu the realizations of certain starred forms are different from what they would be in a simple one-to-one correspondence (there is a formal difference between their surface and deep structures). Thus the very word 'Kikuyu' is 'ɣẹkọyọ' in that language. Similarly,

*-kàté 3 bread > mọ-ɣatɛ 3

*-kíík- put across > -ɣẹẹk- ∅

Compare the following two pairs of realizations:

*-kòtì 5/6, 9/10 nape of neck > i-ɣɔti 5/6

*-kʊ̀ʊ̀t- rub > -ɣʊʊt- ∅

*-kúmì 5/6 ten > i-kọmi 5/6

*-kúnik- cover > -kʊnẹk-

---

1 The occurrence of the second vowel or radical extension vowel, V_e is correlated with the presence of *e in position V_1. See CompB, 1, 1, 44.34.

2 Of two voiceless consonants in two successive syllables the first becomes voiced.

*g > ɣ

E.g. *-gàn- tell a tale > -ɣan-

*-gàndá 9/10 (house) > kẹ-ɣānda 7/8 ø

*-gɨmà DP whole > -ɣima

*-gúím- hunt > -ɣuẹm-

*c > ð

E.g. *-cànɡà 3 sand > mọ-ðāŋga 3

*-cɨ́d- become finished > -ðir-

*-cɨ̀ú 14/6 face > ọ-ðiọ 14/6

*-cùk- wash > -ðuk-

*j > ɸ

E.g. -jí 6 water > ma-ẹ 6

-jó́ yesterday > (ɨ-ɔ 5 day before yesterday)

*-júbà 5 sun > †rẹ-ọa 5; † C.S.2147 *-yúbà

*y > ɸ

E.g. *-yúbà 5 sun > †rẹ-ọa 5; † C.S.955 *-júbà

*-yánà 1/2 child > mw-āna/tʃi-ana 1/8

*-yùb- skin (fruit or animal) > -ọ-

*-yím- (i) become mean (ii) refuse to give > -im-

*m/*n/*ŋ > m/n/ŋ

E.g. *-mèdi- swallow > -mɛri-

*-mɨ̀d- blow nose > -mir-

*-nɨàd- wither > -niar- ø

*-nùà 3/4, 12/13 mouth > ka-nua 12/13

The third correspondence is attested by a single entry in an osculant C.S. :

*-ɲòmbè 9/10 cattle > † .-ɲɔ̄mbɛ 9/10; † *-gòmbè

The following nasal compounds involve mostly Classes 9/10 nominals - a homorganic nasal being the marker of Classes 9/10 independent nominals in Common Bantu.

Thus e.g. '*-pàkò 9/10' read '*mpàkò.'

*mp > .h

E.g. *-pàkò 9/10 tree-hollow > .hakɔ 9/10

*-pɨ́tɨ̀ 9/10 hyena > .hɨtɨ 9/10

*-pʊ̀kò 9/10 mole > .hʊkɔ 9/10

*mb > mb

E.g. *-bàdà 9/10 burn > m-bara 9/10

*-bèdè 9 front > m-bɛrɛ 9

*-bɨ̀tá 9/10 war > m-bɨta 9/10

*-bòdá 11/10 bee-sting > [rọ-βɔra] /m-bɔra 11/10

*-bʊ́ɨ́ 9/10, 11/10 white hair > m-bʊɛ 10

*nt, *nd > nd

E.g. *-tàkà 5, (9) soil > (n-daka 9 mud)

*-ntù 1/2 person > mɔ̄-ndọ 1/2

*-túmbí stool > (n-dūmbɛ 9/10 wooden head rest)

*-dèdʊ̀ 7/8. 9/10, 10, 11/10 beard > n-dɛrʊ 9/10

*-dùgú (9/10) relative > n-dɔrọ 9/10

*-dòndà 7/8 sore > kɛ- rɔ̄nda 7/8

*ŋk, *ŋg > ŋg

E.g. *-káŋgà 5/6, 9/10 guinea fowl > ŋ-gāŋga 9/10

*-kíɣmà 9 bread; mush > ŋ-gima 9

*-kúdù 9/10 tortoise > ŋ-guru 9/10

*-nùŋk- smell > -nũŋg-

*-ŋgí DP other > -ŋgɛ

*-gì 9/10. 11/10 fly > ŋgi 9/10

*-gùbú 9/6. 9/10 hippopotamus > ŋ-guɔ 9/10

*nc > .ð

E.g. *-cá kindness > .ða 9

*-cùdì wind per anum > .ðuri 9

*-cóni 9/ 10/ 11 shame > .ðɔni 9

*-cí 9 ground > .ðɛ 9

*nj > nɟ

E.g. *-jìdá 9/10 path > n-ɟera 9/10

*-jògù 9/10 elephant > n-ɟɔɣu 9/10

*-júídí (9/10), 11/10 hair > n-ɟuɛrɛ 9/10

2.5.3 Skewed reflexes of Common Bantu in Kikuyu

As compared with the other Selected Languages the Kikuyu data reveal a remarkably small number of skewed reflexes. See table on p.51.

2.5.3.1 Reflexes skewed for meaning

With respect to the meaning, 37 skewed reflexes were counted. Since a detailed examination of the category of meaning does not

directly bear upon our investigation we shall not dwell on it. Suffice it to mention that a wide gamut exists of various degrees of departure from the connecting meaning.[1] The following are examples of this. Asterisk before an expression indicates a starred form meaning, the expression right of the arrow is the meaning in Kikuyu.

* thicket →(bush country)

* pull →(take out)

* millet, eleusine →(maize)

* kind of rat →(hare)

The examples are all taken from entries whose shape corresponds perfectly with their starred forms.

2.5.3.2 Reflexes skewed for shape

Twenty-six items were counted whose shape diverged from the regular correspondence pattern. This, it is worth repeating, is a *very* low figure dealing as we are with a Test Language, where the body of data originally screened against Common Bantu was of a considerable size.

The only clearly identifiable tendency observed among the items skewed for shape was the large proportion of vocalic skewing. In fact it accounted for a half of the skewings. Compared with the other Selected Languages this is a large proportion, but since Kikuyu, as a seven-vowel language, has a larger number of potential vocalic reflexes, direct or skewed, this is only a logical result.

The following are examples of vocalic skewing in Kikuyu.

---

1 As indeed in the remaining Selected Languages and in Bantu languages in general.

*-kóbà 3/4 strap > [mǫ-kǫa 3/4; * → -kɲa]

*-píù DP hot > [-hiǫ; * → hǫǫ]

*-yújì 11/10 river > [rǫ-ǫǫ 11/10; * → -ǫi]

The only skewing which occurred more than once involved β, a sound extraneous to Kikuyu. It appeared in the following skewed correspondences:

*-bààbà 1a/2 father > [βaaβa 1a; *b → ɸ]

*-bàdù 11/10 rib > [rǫ-βaru 11; *b → ɸ]

*-bòdá 11/10 bee-sting > [rǫ-βɔra] /m-bɔra 11/10

*-cúpà 9/10 jar > [tʃɥɥβa 9/10; tʃɥ,β ≠; *nc → .ʤ, *p → h]

*-kóópì 5/6 slap > [i-kɔɔβi; *p → h]

The last two Kikuyu entries are loan words from Swahili, since Swahili p corresponds to Kikuyu β in Swahili loans. β is extraneous to Kikuyu, indicating thus alien material. And although similar alien material is present in the first three entries, namely the extraneous β, the likelihood of contamination from Swahili is remote. For an assessment of the situation it is helpful to turn for evidence to the actual composition of Guthrie's C.S. and to the accompanying commentary.

C.S.7 *-bààbà 1a/2 'father,' contains only four direct reflexes.[1] There are also appended one item skewed for meaning, and as many as eight items skewed for shape, of which the Kikuyu is one. The sizeable commentary to this C.S. concludes that 'the prehistory both of this C.S. and of the skewed items in the footnotes is uncertain'.[2]

---

1 CompB, II, 3, p.18.

2 Ibid.

The second item, roβaru, is associated with C.S.30, the evidence of which is similarly inconclusive from the point of view of whether it might be a loan from Swahili. C.S.30 is a series of a substantial size with twenty entries. There is only a single skewed item appended to it, namely the Kikuyu item in question. Thus it is likely that we are dealing here with a C.S. the source-item of which existed in Proto-Bantu period.[1] The Kikuyu item, skewed in one 'minor' feature, and in the same class as the starred form, may have been contaminated sometime in the Bantu period. But whether the contamination came from Swahili or from another source is uncertain.

Although the third item, /roβɔra// mbɔra, is skewed in its singular form, the fact that its plural is a valid entry in a C.S. makes the item as a whole a direct reflex.

The following is a tabular summary of the Kikuyu skewed reflexes of Common Bantu.

2.5.3.3 Numerical summary

SKEWED REFLEXES OF C.B. IN KIKUYU

Meaning skewed		35
Shape skewed		26
Vowels skewed	9	
Consonants skewed	12	
Both skewed	5	
Both skewed		2
Total		63

1 CompB, II, 3, p.23.

2.5.4 Conclusions

As compared with the other Selected Languages and, indeed, with the other Test Languages, Kikuyu was not found to contain a very high number of Common Bantu reflexes. This fact is not to be attributed to a limitation in the size of the original data since Kikuyu is one of Guthrie's Test Languages, which were subject of a detailed examination on the basis of an extensive range of data. It rather correlates with Guthrie's statement that the number of C.B. reflexes decreases with distance from his Proto-Bantu nucleus area.[1]

An overall characteristic of the Kikuyu sound-correspondences was their straightforwardness, which may be owing to the Kikuyu seven vowel qualities which correspond one-to-one to the seven starred vowels of Common Bantu. Thus for example the starred *d, which is attested by a great many items in all the six Selected Languages, corresponded with the utmost regularity to the Kikuyu r. A similar d-sound correspondence is absent from the other Selected Languages, in which *d presents a complex type of sound-correspondence.

Apart from the 414 direct reflexes there were 63 skewed items counted among the data, representing 15.2% of the direct reflexes: this is an unusually low figure in the context of the Selected Languages. Equally unusual among the Selected Languages is the very small number of reflexes skewed for shape. Among 63 skewed items there were only 26 skewed for shape.

---

1 The Proto-Bantu nucleus area was located approximately in the region of present-day Katanga.

2.6 REFLEXES OF COMMON BANTU IN MVITA AND UNGUJA

2.6.1 General

The two Swahili dialects, G.42b and G.42d, display mostly identical sound-correspondences with C.B., both in open- and close-vowel environments. Different developments affect only two consonants and one nasal compound, *b, *c, and *nj, and even those not in all phonetic environments.

Correspondences are very frequently attested by reflexes occurring jointly in the two languages: indeed more reflexes are shared by Mvita and Unguja than by any other two Selected Languages. This will be discussed further in chapter 3.

An added reason why the two Swahili dialects will be discussed simultaneously in this section is that it will make it possible for immediate comparisons to be made between Mvita and the far more familiar Unguja.

Unguja material in CompB is extensive since G.42d is a Test Language. However, it has been revised as a test case of Guthrie's treatment of CompB data, [1] and those items which have been added or corrected will be indicated in the text by a raised plus sign.

Because of its close relationship to Unguja, only the very limited number of fifteen entries are quoted from Mvita in CompB, namely those with distinct correspondences for *c and *nj. Therefore all data quoted from Mvita are samples

---

1 Published jointly with M.A.Bryan as 'Comparative Bantu: the case of two Swahili dialects' in ALS, XIV, 1973. Comparison with CompB is not an aim of this study, even though some important differences have to be noted. The present author was responsible for the Mvita data.

2 The statement of the Mvita sound-shifts in CompB, I, 2, p.50, (cont. ...)

of this author's data, unless otherwise specified.

The following table demonstrates the positions of Mvita and Unguja among the Test Languages, according to the number of C.B. reflexes each of them contains.

1	Bemba M.42	800 reflexes of C.B.
2	UNGUJA G.42d	674 [1]
3	Kongo H.16	618
4	Luba-Katanga L.33	606
	MVITA G.42b	580
5	Yao P.21	579
6	Luba-Lulua L.31	557
7	Sukuma F.21	547
⋮		⋮
28	Duala A.24	215

(..(cont.)

is as follows: 'Broadly similar to G.42d, but *c > ʧ; *j/*nj > ʤ/nd'. Guthrie fails to note the different correspondence of *ɓ, and the fact that *j has the same reflex as in Unguja.

1 M.Mann's figure, 656, is based on the Register of entries from the Test Languages, CompB, I, 2, 65-104.

Consonant reflexes of Common Bantu in Mvita and Unguja.

Mvita reflexes given before, Unguja after colon.
(A blank in a column indicates no or inconclusive evidence.)

	+*-a[1]	+ +*-$\underline{i}$	+*-$\underline{u}$	*-$\underline{i}$+V[2]	+*-$\underline{u}$+V[2]	*-i+V[2]
*p	p,p	f,f	f,f	fy,s	f,f	p,p
*b	ʋ,w[3]	v,v	v,v	vy,z		
*t	t,t	s,s	f,-	s,s		
*d	ɸ,ɸ[4]	z,z	v,v	z,z	v,v	l,l
*k	k,k	ʃ,ʃ	f,f	ʃ,ʃ	f,f	tʃ,tʃ
*g	ɠ,ɠ	z,z	v,v	ʒ,ʒ	v,v	
*c	ʨ,tʃ	s,s	ʨ,s	s,s		
*j	ɟ,ɟ	ʒ,ʒ	ɟ,ʒ			
*y_1	ɸ,ɸ5	ɸ,ɸ	ɸ,ɸ			
*mp	ph,ph	f,f	f,f			
*mb	mb,mb	mv,mv	mv,mv			
*nt	th,th	s,s	f,-			
*nd	nd,nd	nz,nz				
*ŋk	kh,kh	ʃ,ʃ	f,f			
*ŋg	ŋg,ŋg	nz,nz				
*nc	ʨh,tʃh	s,s				
*nj	ɲɟ,ndʒ					

1 *a or any non-cedilla vowel

2 Any vowel other than preceding.

3 ɸ and ɓ are the reflexes of *b in specific contexts, which are discussed under 2.6.2.

4 l is a direct reflex in the context *d +(*e/*i).

5 *y_2 has as its reflex a palatal semivowel [j].

2.6.2 Direct reflexes of Common Bantu in Mvita and Unguja

The same order of presentation will be followed here as in the previous sections: individual correspondences will be given and exemplified in the first part. Where a correspondence was postulated on the evidence of less than three items, it will be quoted in full. In the second part of the section we shall introduce the skewed items and any existing tendencies among them.

Identical correspondences between C.B. and Mvita and Unguja will not be specially indicated. Thus *p > p means that the correspondence is valid for both languages. Where two different correspondences were postulated, these will be quoted.

Simple correspondences, like the one following, will be illustrated by three examples of *Ca, *Ce/*-i and *Co/*-u for open-vowel environment, and one each from close-vowel environment.

2.6.2.1 Consonant reflexes in open-vowel environment

*p > p

The attestation of this simple correspondence by over sixty items implies that *p is one of the most productive starred consonants in Swahili. Most attesting items occur jointly in the two languages. E.g.

*-pápati$_{s}$- touch; grope > -papas-

*-pépò 9/10, 11/10 wind > u-pepo 11/10

*-pònd- pound > -pond-

*b > Mv. ϕ , U. w

b + (-o/*-u) > ϕ

*b + (*u-*o/*u)nominals) > ɓ

For rules generating ɓ see also 2.6.2.2, *ib$_s$ and *yib$_s$.

*b is the most complex starred consonant in Swahili and although the total number of attesting items is by no means small, the correspondences postulated for restricted environments are not unequivocal.

The first and second correspondences are attested by 48 and 13 items respectively, the third one by only four.

E.g. *-bá become, be > Mv. -ϕ-, U. -w-

*-bé you sing > Mv. ϕeϕe Rd, U. wewe Rd

*-bìmbì 5/6 wave > Mv. ϕimbi 5/6, U. wimbi 5/6

*-bòd- become rotten > -oz- [ɣ]

*-bón- see > -on-

*-búd- kill > -u-

*-búnò 7/8 waist > ki-uno 7/8

*-bògà 9/10 vegetable > (m̄-boɠa 3/4 pumpkin plant)

*-bónò (i) 3/4 castor-oil plant > m̄-ɓono 3/4

*-bòŋgó 14 brain > U. u-ɓoŋgo 14; Mv. entry is in Cl.5 and subject to a different correspondence

*-bùyú 3/4 baobab > m̄-ɓuyu 3/4

*t > t

Another abundantly attested correspondence, this time with over ninety attesting items.

E.g. *-tátud- untie > (-tatu- untangle)

*-tátù DP three > -tatu

*-té 5, 6, 11 spittle > +ma-te 6

*-tí 3/4, 7/8 tree > m̄-ti 3/4

*-túm- send > -tum- spec

*d > ø; *d + (*-e/*-i) > l

E.g. *-dàp- swear > -ap-

*-dómb- ask for > -omb-

*-dúmik- bleed by cupping > -umik-

*-dèèdó today > leo

*-dáádik- invite > +-alik-

*-dìmìdà 7 Pleiades > ki-limia 7

*k > k; *k (*-i)+V > tʃV

The first correspondence is attested by well over one hundred items, the second by three items. This (latter) also affects developments in Cl.7 independent prefix of vocalic stems; e.g. *ki- + *-ùmbá > tʃ-umba.

E.g. *-kàdé olden times > kale

*-kídà 3/4 (tail) > m̄-kia 3/4

*-kómbè 5/6 shoulderblade > kombe 5/6

*-kí- dawn > -tʃ-

*g > ɠ

Attested by over forty items, e.g.

*-gàb- divide > Mv. -ɠavany-, U. -ɠaw-

*-gègò 5/6 (ii) molar tooth > ɠeɠo 5/6

*-gùdù 3/4, 5/6, 15/6 leg > Mv. ɠuu 5/6, U. m̄-ɠuu 3/4

*c > Mv. ʈ, U. tʃ

Attested by over sixty items, e.g.

*-càn- comb > Mv. -ʈan-, U. -tʃan-

*-cínj- butcher > Mv. -ʈiɲɟ-, U. -tʃindʒ-

*-cùdʒ 3 broth > Mv. m̄-ʈuzi 3/4, U. m̄-tʃuzi 3/4

*j > ʒ

Attested by twelve items, e.g.

*-jàdà 5/6 rubbish-heap > ʒaa 5/6

*-jèŋg- build > -ʒeŋg-

-jí/-jí$_s$ 6 water > ma-ʒi 6

*-júúdì$_s$ 5 day before yesterday > ʒuzi 5

*y_1 > ø; *y_2 > j [palatal semivowel]

First correspondence attested by over seventy items, second by nine items. E.g.

*-yényé 1/2 (i) owner > mw-enye 1/2

*-yíg- imitate > Mv. -igiz-, U. -ig-

*-yòyá 5/6, 11/10 feather > Mv. nyoya 10, U. nyoya 5/6 NA, +unyoya/nyoya 11/10 NA

*-yòyò 3/4 heart > m-oyo 3/4; +nyoyo 10 NA spec

*-búy- come (or go) back > +-uy-

The nasal consonants *m, *n, *ŋ displayed in both languages simple correspondences to m, n, ŋ respectively. Although they were not attested by as many items as the preceding correspondences, the attestation is ample.

E.g. *-mèd- grow, sprout > -me-

*-néén- speak > -nen-

*-nàŋ- grumble > U.(-ŋaŋany- pester)

*-ŋòmbè 9/10 cattle > ŋombe 9/10

The last item's validity is subject to the operation of Meinhof's rule in Swahili.[1]

---

1 A nasal sequence rule, it was first formulated by Meinhof for Ganda - hence its alternative term 'the Ganda Law'. In Swahili it operates partially. It affects nasal compounds in first position followed by another nasal compound in second position so that they correspond to the appropriate single nasal, e.g. *-dùŋgú 9/10 'porcupine' > nungu 9/10.

The nasal compounds display a well-defined tendency among their reflexes. When the second component is voiceless, the compound corresponds to that second component aspirated, e.g. *mp > ph. When second component is voiced, the compound has a simple one-to-one correspondence, e.g. *mb > mb. Implosion in g disappears in the velar nasal compound, and *nj > Mv. nd , U. ndʒ.

Each nasal compound will be illustrated by one example.

*-pándà 5/6, 9/10 fork (in road, etc.) > phanda 9/10

*-bòɠà 9/10 vegetable > m-boɠa 9/10

*-ntù 1/2 person > m̄-thu 1/2

*-dèɠè 9/10 (i) bird > n-deɠe 9/10

*-kúápà 5/6, 9/10 armpit > Mv. khwapa 9/10,
U. kwapa 5/6

*-kátà 9/10 headpad > khata 9/10

*-gòmà 9/6, 9/10 drum > ŋ-goma 9/10

*-cúpà 9/10 jar > Mv. thupa 9/10, U. tʃhupa 9/10

*-jàdà 9 (i) hunger (ii) famine > Mv. ɲ-ɟaa 9/10,
U. n-dʒaa 9/10

2.6.2.2 Consonant reflexes in close-vowel environment

There are even fewer differences between reflexes in Mvita and Unguja in close-vowel environment than there are in open-vowel environment. In the nature of the case fewer attestations are available, but in the less restricted context of *-i̯ and *-u̯ following single consonants the attestation is usually sufficient.

Below, one example each will be quoted of correspondences with sufficient evidence to support them. All

attesting items will be quoted when they are less than three.

p + (-i/*-u) > f; *p(*-i)+V > Mv. fy, U. s

E.g. *-pic- hide > Mv. -fit-, U. -fitʃ-

*-pako 3/4 bag > m̄-fuko 3/4

*-piagid- sweep > Mv. -fyagi-

*-piat- seize > Mv. (-fyat- contract, pull in)

*-piom- read > U. -som-

*-piu 7/8 knife > U. ki-su 7/8

Although the last four items are the only evidence on which the correspondences were postulated, a parallel correspondence affecting *b might be accepted as corroborative evidence.

b + (-i, *-u) > v; *b (*-i)+V > Mv. vy, U. z;

*b + (*i-) > ɓ

E.g. *-bi 6, 13 excreta > ma-vi 6

*-bun- harvest > -vun-

*-biad- bear (child) > Mv. -vya-, U. -z-

*-biada 1/2 cousin > Mv.(mavyaa 1/2 maA mother-in-law of wife)

*-bega 5/6 shoulder > ɓega 5/6 YA

*-biya 7/8, 9/10 pot > U. ɓia 5/6 YA

*-dib- stop up > -ziɓ-

*-dibo (7/8) stopper > +ki-ziɓo 7/8

t + (-i) > s; *t + (*-u) > f; *t(*-i)+V > s

E.g. *-tiki 7/8 stump of tree > ki-siki 7/8

*-tumo 5/6 spear > fumo 5/6

*-tian- forge > U. -san-

The last item above provides the only attestation in Unguja of the postulated correspondence. It is absent in Mvita.

d +(-i$_S$) > z; *d +(*-u$_S$) > v

Identical correspondences operate when *d is followed by a close vowel and another vowel, but the first vowel has no surface reflex.

E.g. *-dígò 3/4 load > m̄-zigo 3/4

*-dìàm- sink > -zam-

*-dùm- roar, rumble > -vum-

*-dúád- wear > -va-

*-dúík- clothe > -vik-

k + (-i$_S$) > ʃ ; *k + (*-u$_S$) > f

As in the preceding *d correspondences, the sequence *CiV$_S$ has as its reflex CV. See the following examples.

*-kínà 5/6 base of tree trunk > ʃina 5/6

*-bàki- light fire > Mv. -βaʃ-, U. -waʃ-

*-kútà 6 oil > ma-futa 6

*-kúò 7, (11) death > ki-fo 7/$^+$8

g + (-i$_S$) > z; *g + (*-u$_S$) > v; *g (*-i$_S$)+V > ʒ

E.g. *-gìmà DP whole > -zima

*-gùnd- become rotten > -vund-

*-gìà 1/2 slave > m̄-ʒa 1/2;(m̄-ʒakazi 1/2 female slave)

The last item is the only attestation of that correspondence.

c + (-i/*-u$_S$) > s; *c + (*-u$_S$) > Mv. t

While the *ci$_S$ correspondence is amply attested by eight items, the *cu$_S$ correspondence is based on the evidence of a single item in each language.

*-címá 7/8 > ki-sima 7/8; + *-tímá

*-cìó 14/6 face > u-so 11/10

*-cùg- wash > U.(-sugu- rub)

*-cùng- look after > Mv. -tung-

j + (-i$_S$/*-u$_S$) > ʒ

A very tentative correspondence, which was based on the following evidence:

-jí/-jí$_S$ 6 water > ma-ʒi 6

*-júb- know > -ʒu-

*yi$_S$ > ø; *yib > iɓ [1]

E.g. *-yí$_S$cò 5/6 eye > Mv. ʒi-ʧo 5/6, U. ʒi-ʧo 5/6

*-yí$_S$b- steal > -iɓ-

*-yí$_S$bà 3/4 thorn > mw-iɓa 3/4

*-yí$_S$bì$_S$ 1/2 thief > Mv. mw-ivi/vevi 1/2,

U. mw-ivi/wevi 1/2

*-yí$_S$buk- come out of the water > U. -iɓuk-

Although the evidence has been rejected as valid elsewhere [2], we shall accept them tentatively on account of the fact that apart from the above entries no skewed item was elicited for the correspondence. Together with the correspondence *ib$_S$ > ɓ they form mutually corroborative evidence.

Nasal compounds in close-vowel environment represent a most restricted phonetic environment, with a number of correspondences established on less than three items. As far as evidence was available it can be concluded that the voiceless compounds correspond to the reflex of their second component, e.g. *mp + (*-i$_S$/*-u$_S$) > f.

The voiced nasal compounds correspond to the appropriate nasal and the reflex of its second component, e.g. *ndi$_S$ > nzi.

---

1 See also *ib$_S$ > ɓ.

2 Magdalena Slavíková and Margaret M.Bryan, op. cit.

The following correspondences have been established :

mp + (-i/*-u) > f

*-pígò 9/10 kidney > Mv. figo 9/10

*-pímbò 9/10 stick > fimbo 9/10

*-pìtì 9/10 hyena > fisi 9/10

*-púngàté 9 seven > Mv. fungate 9 , (fungate 9 seven days' honeymoon or seven days' mourning for infant)

U.(fungate +9 seven days' honeymoon)

mb + (-i/*-u) > mv

Evidence for the first part of the correspondence appears to be inconclusive but, if the second correspondence is accepted as corroborative evidence, it may be tentatively accepted as valid, and vita rejected as skewed.

-bímb- swell/-bímbò (7/8) swelling > (mimba 9 pregnancy)[1]

*-bìtá 5/6, 9/10 war > Mv. [vita 9 or 8], U. [vita 9 or 8]

-búì/-búì 9/10, 11/10 white hair > m̈-vi 9/10

*-búdà 9 rain > Mv. m-vua 9/10, U. m-vua 9

nt + (-i) > s; *nt + (*-u) > f

-tíkù/-cíkù (5/6), 9/10 day of 24 hours > siku 9/10

*-tíkù 14/6 night > u-siku 14/6, U. also 11/10

*-tùngò 9/10 civet-cat > fungo 9/10

---

1 Due to the operation of 'Meinhof's rule' of nasal sequence.

nd + (-i/*-u) > nz

*-tund- teach > -funz- [Y]

*-cándú 5/6 branch > Mv. ki-tanzu 7/8, U. [tanzu 5/6 ⊙]

*nk + *-i > ʃ ; *nk + (*-u) > f

*-kíngò (i) 9/10 neck; nape (ii) 9/10 neck > Mv. (i) ʃingo 5/6, U. (ii) ʃingo 9/10

*-kúmù 1a/2, 9/10 chief > Mv. [fumo 9/10⊙]

*-kúpí (i) DP short > -fupi; DP Cl.9/10 = zero.

ng + (-i) > nz; *ngu: no evidence available

*-gì 9/10, 11/10 fly > n̄-zi 9/10

*-gìgè 9/10 locust > n-zige 9/10

*-gìmà DP whole > -zima; DP Cl.9/10 = nz-.

nc + (-i) > s; *ncu: no evidence available

E.g. *-címbà 5/6, 9/10 lion > simba 9/10

No evidence was elicited for *nji, *nju correspondences.

Finally, aspiration, which plays a role in both languages, displays the following recurrent pattern : when expected on the final consonant, it is transferred to a preceding plosive. It may appear either as a free variant, as in *-ntù 7/8 'thing' > kithu 7/8, also +khitu 7/8, or as the only reflex, as in *-tùnt- throb > Mv.(-thut- pant).

2.6.3 Skewed reflexes of Common Bantu in Mvita and Unguja

More skewed reflexes were identified in Mvita than in Unguja - its proportion of skewed to direct cognates was by 4% larger. A table at the end of this section compares the results.

Mvita yielded altogether 166 skewed reflexes, Unguja 160. Both had a relatively high proportion of items with skewed meaning: G.42b had 62, G.42d 71. Among the remaining Selected Languages only Kikuyu had a similarly high proportion.[1] Both languages further contained a certain number of double entries in C.S., one of which was valid, one was not. E.g. *-bídu- 'become cooked' > Mv. [-iv-], U. -wiv-, +[-iv-], a fact which was not to be observed in the remaining Selected Languages.

In general it appeared that there were few substantial or multiple skewings, and very few skewings were vocalic. In this respect Saghala is prominent among the Selected Languages with its substantial number of 'major' and multiple skewings.

The following table summarizes the results. Since, as among all the Selected Languages, and among Bantu languages in general, verbal radicals prevail over nominal stems, a single figure expresses the total.

SKEWED REFLEXES OF C.B. IN MVITA AND UNGUJA

Mvita G.42b	[96]	(62)	[(8)]	T. 166
Unguja G.42d	[78]	(71)	[(11)]	T. 160

2.6.3.1 Nature of individual skewings

*p

Among the large collection of over sixty attestations there

1 In fact Kikuyu had more items with skewed meaning than with skewed shape.

were a mere four departures encountered from the regular reflex p. Three of them occurred in the *piV pattern with different results in Mvita and Unguja. (Mvita entry accepted as valid.)

*-pi$_{s}$ágid- sweep > U. ⟨-fa$_{s}$i- ⟵ *-pu$_{s}$ágid-⟩

The other two skewings are of the type *pi$_{s}$V > fyV in Unguja items. The 'regular' Unguja reflex s was postulated on two attesting items together with comparative evidence of *bi$_{s}$V > zV.

In Mvita the situation is reverse, with s as the skewed reflex. Hence,

*-pi$_{s}$ómo- read > Mv.⟨-som-⟩

*-pi$_{s}$ú 7/8 knife > Mv. ⟨ki-su 7/8⟩

It is interesting to note that the two reflexes coexist in an Unguja entry, -som- 'read' and ⟨-fyom-⟩ 'recite Koran'.

*b

The fact that three different direct reflexes exist each in a particular environment, has given rise to a situation in which the three reflexes, ʋ/w, ɸ and ɓ, intermingle and appear in 'wrong' phonetic contexts.

No 'substantial' skewing was noted.

*t

A single consonantal skewing was noted among the exceptionally high number of ninety attesting items. ʃi appeared where si would normally be expected as a reflex of *ti$_{s}$. But the Unguja entry in question contained both a direct and the skewed item just mentioned.

*d

A situation similar to that under *b was to be observed here. Two direct reflexes exist, each in a particular phonetic context. l ⟨ *d + (*-e/*-i), zero occurs in the remaining open-vowel environments. The skewings contain the two reflexes interchanged in their respective environments. There is also r, which occurred in either of the two contexts, and ɗ, both of which are extraneous to Swahili.

*k, *g

One skewing was noted of the first consonant, namely *ki > si instead of the 'regular' ʃi.

Only three skewings were identified of the second consonant. Two of them, in Mvita, had zero as a reflex of *g , the third, occurring jointly in both languages, had ʒ:

*-góŋgòdò 5/6 millipede > [ʒoŋgoo 5/6]

*c

Since the two languages have different reflexes of *c, it was possible to identify their influence on each other by detecting their mutual loan words. Unguja's in Mvita contained tʃ while Mvita's in Unguja contained an alveolar t (there is no dental variety in Unguja's sound inventory).

There were several instances of ʃ occurring in both languages instead of the regular t or tʃ in Mvita and Unguja respectively. ʃ also occurred in both languages as a skewed reflex of *c in open- and close-vowel context.

E.g. *-cáŋgá 14 bead > [u-ʃaŋga 11 or 14/10]

*-còkó 9/10 kind of monkey > Mv. [ʃokwe 5/6],

U. [ʃokwe 9/10 chimpanzee]

*-cíd- become finished > Mv. [-ʃ-]

*-còn- sew > [-ʃon-]

The single nasal consonants and the nasal compounds yielded remarkably few skewed reflexes. The only recurring items involved *nj, which has a distinct reflex in either language. According to their interchange it was possible to identify loan words between the two languages, e.g

*-jògù 9/10 elephant > Mv. n-ɟovu 9/10,

U. [n-dovu; * → ndʒovu],

in which the Unguja item is probably a loan from Mvita.

*nc yielded three skewed reflexes, all identical. s occurred in items where tʃh would be expected in Unguja and ʧh in Mvita, as in

*-cóni 9, 10, 11 shame > +[soni 9@]

2.6.3.2 Numerical summary

Proportion of skewed to direct reflexes is summarized in the following table.

	Mvita G.42b	Unguja G.42d
Skewed reflexes	166	160
Skewed shape	96	78
Skewed meaning	62	71
Both skewed	8	11
Direct reflexes	580	674
Proportion of skewed to direct reflexes	29%	25%

3 SELECTED LANGUAGES, AS COMPONENTS OF COMMON BANTU, COMPARED AMONG THEMSELVES

3.1 GENERAL

Having established in the preceding chapter the sound-correspondences between Common Bantu and each Selected Language and, on the basis of these, having determined the Common Bantu reflexes in each of them, we shall now proceed to employ these 'primary' results in certain computations in order to formulate finally hypotheses of relationships among the Selected Languages.

From among the six languages the following combinations of pairs were established so that each language would be compared with every other language. The total number of pairs to be investigated can be presented in a 'Triangle of Relationships'.[1]

D	S	G	K	M	U
DS	DG	DK	DM	DU	
	SG	SK	SM	SU	
		GK	GM	GU	
			KM	KU	
				MU	

The 'Triangle of Relationships', representing all feasible combinations in pairs of the following six languages: D=Dawida, S=Saghala, G=Giryama, K=Kikuyu, M=Mvita, U=Unguja.

---

[1] M.Mann points out the elementary nature of this concept from a statistician's point of view.

A number of indices have been obtained for each language, based on both the numbers of direct and skewed reflexes. The present study being concerned with the Common Bantu element as a measure of relationship between these languages, we shall not go beyond presenting the figures for the proportions of skewed reflexes in the Selected Languages, while the direct reflexes will be discussed at length.

The 'skewed figures' are not excluded altogether as they are an important complementary element to the study, and present a body of hitherto unknown data.

Since certain formal statistical methods were used to express the degrees of relationship a brief explanation will follow of the processes involved.

3.1.1 The Index of Relationship

Guthrie's formula was employed in this work to express the measure of closeness of relationship between languages. Other measures of similarity (or dissimilarity) have been used elsewhere[1], and a mention will be made of two of them and their relation to the IR further on.

The mathematical process involved in calculating the Index of Relationship (IR) has been described before[2], but a brief explanation of the composition and function of the formula is helpful at this point.

---

1 One of them is Ellegard's r_n, the other Linguistic Distance, D, both used in numerical taxonomy. Guthrie's formula, $IR = \left(\frac{1000q}{x+y+q}\right) / \left(\frac{x}{x+q} + \frac{y}{y+q}\right)$, has been criticised for being unnecessarily complex and having 'the awkward property of tending to infinity as languages become closer'.
See A.Henrici, 'Numerical classification of Bantu languages', ALS, XIV, 1973, 82-104.

2 CompB, I, 1, Appendix 6/3; also M.Guthrie, 'Some uses of arithmetical computation in comparative Bantu studies', Transactions of the Philological Society, 1964, 108-28; also R.Jones-Phillipson, op.cit.

Three essential characteristics were combined in the formula which is a measure of similarity:

q - the total number of reflexes occurring jointly in a pair of languages,

x - the total number of reflexes exclusive to the first language,

y - the total number of reflexes exclusive to the second language.

These three characteristics are combined into the two main components of the IR, which may be termed Q and C.

$Q = \frac{q}{x+y+q}$, i.e. the ratio of common reflexes to the total number of reflexes represented in either language.

Q increases as the number of common reflexes increases, and decreases as the number of exclusive[1] reflexes increases. It is not affected by the relative value of x and y.

$C = \frac{1}{(\frac{x}{x+q} + \frac{y}{y+q})}$, i.e. C is inversely proportional to the sum of the ratios of the exclusive reflexes to the total number of reflexes in each language.

C increases as the difference between x and y increases. The greater the discrepancy between the number of reflexes in one language and the number of reflexes in the other, the greater the value of C.

C also incidentally increases as the number of common reflexes increases, and decreases as the total number of exclusive reflexes increases thereby exaggerating the range of the final index.

C, or Guthrie's correction of the basic proportion of reflexes for the two languages, Q, is justified by the greater significance of a high proportion of common reflexes in an innovating as opposed to a conservative language. Accordingly, the index is weighted towards the innovating language of the two.

---

[1] Other synonyms are used in this work, too: 'unique' and 'peculiar'.

Although not immediately apparent from the following table of the Hierarchy of Closeness of Relationship among the Selected Languages, pairs 14 and 15 can serve as a demonstration of this. The former pair has less joint reflexes than the latter, 165 against 170, but the large proportion of unique reflexes in one of its members increases the resulting IR as compared with the latter pair. This, pair 15, would merit at a first glance a higher IR value than no.14: it has less unique reflexes and more joint reflexes, both pointing towards a higher IR. The difference is in the relatively more balanced proportions of unique reflexes in each language.

For further exemplification two sets of data were taken to demonstrate the difference in IR values caused by the discrepancy between the numbers of unique reflexes in each language, x and y.

x = 320	x = 133
y = 320	y = 507
q = 165	q = 165
IR = 155	IR = 178

Note that the total value of x+y is the same in both cases.

To sum up the interpretation of the IR formula and its constituents,

- the higher the IR, the closer the relationship;
- the larger the number of jointly occurring reflexes, the closer the relationship;

Provided that q is greater than zero:

- the smaller the number of reflexes unique to each language, the closer the relationship;
- the greater the discrepancy between the number of reflexes in one language and the number of reflexes in the other, the closer the relationship.

The following table summarizes numerically the nature of the hierarchy of closeness of relationship among the six Selected Languages. The figures which are pertinent to this study appear in the first four columns.

As stated in the previous chapter, all figures are this author's own, and based on her own data except for Kikuyu E.51, the data for which were extracted from the Catalogue of Common Bantu, CompB, II, 3 and 4. Although care was taken in the process of counting the individual entries, a degree of fluctuation will be noted among the figures. It ranges between 1 - 5% among the direct reflex figures and slightly over 5% among the skewed figures, and is largely due to uncertain judgement of the validity of some entries.

The Hierarchy of Closeness of Relationship.

M=Mvita U=Unguja G=Giryama
D=Dawida S=Saghala K=Kikuyu

	Languages investig-ated:X+Y	IR	Direct reflexes: Joint reflexes (q)	Direct reflexes: Reflexes unique to X	Direct reflexes: Reflexes unique to Y	Skewed reflexes: Joint skewed refl. (q)	Skewed reflexes: Skewed unique to X	Skewed reflexes: Skewed unique to Y
1	M+U	2089	507	56	158	111	80	59
2	G+M	737	333	128	214	18	88	159
3	D+S	642	188	105	110	43	86	79
4	G+U	612	348	112	324	18	90	144
5	S+G	364	196	109	270	17	112	93
6	D+G	294	172	116	283	16	115	95
7	S+M	247	180	115	385	12	106	169
8	D+K	221	141	146	263	10	123	56
9	G+K	211	168	274	225	7	98	60
10	S+U	207	181	114	487	17	106	152
11	K+U	198	205	205	458	6	58	163
12	S+K	185	130	168	269	9	112	58
13	D+M	177	151	142	408	11	115	173
14	D+U	171	165	133	506	10	120	157
15	K+M	170	170	228	390	9	53	173

3.1.2 Other indices of relationship

Apart from Guthrie's formula for calculating the IR, other indices are available which indicate certain types of linguistic relationship. They had been arrived at in the process of calculating the IR.

a) Reflexes occurring jointly

One of the essential indices is that expressing the number of reflexes shared by the languages in question. Figures for the Selected Languages appear in column 'Joint reflexes(q)' on previous page.

However important a marker, it alone would yield a substantially different account of the relationships of languages from one which takes into account the unique reflexes.

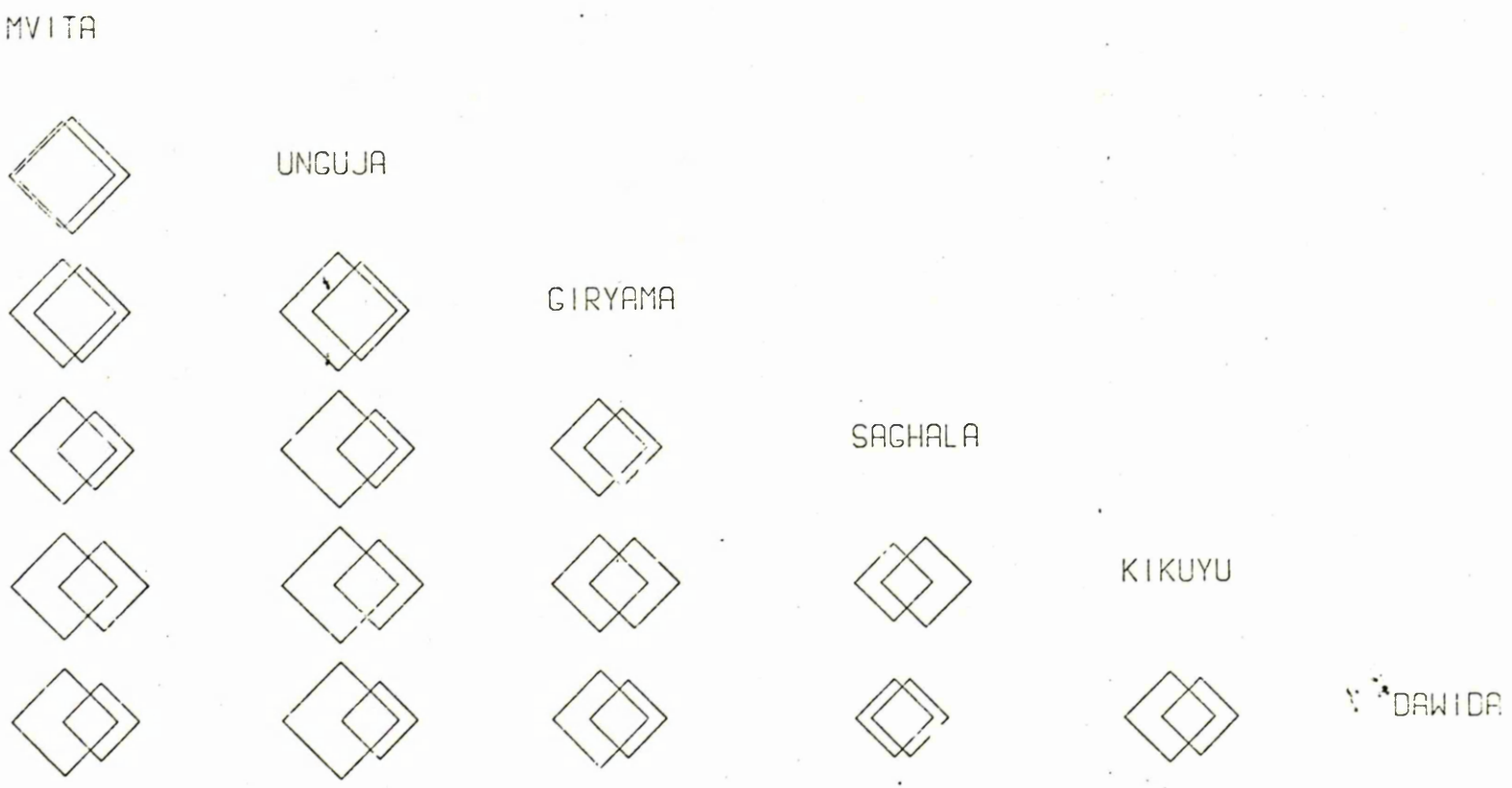
Nonetheless, such supplementary criteria need not always alter the conclusions reached on the evidence of the joint reflexes only.

Mvita and Unguja share the highest number of reflexes, 507, as well as the least number of reflexes unique to each, 56 and 158 respectively. The value of their IR is the highest, too, - 2089.[1]

On the other hand the case of Kikuyu and Mvita is far from uncommon. Notwithstanding the fact that their relationship stands at the very end of the scale opposite to Mvita and Unguja, Kikuyu and Mvita share not a small number of reflexes; 170. The reflexes unique to each of them outnumber by far this figure, they are 228 and 390. These latter two necessarily bear upon the first, thus producing the lowest IR figure among the fifteen pairs.

---

[1] It is of interest to compare the index figure obtained for Mvita and Unguja, G.42b and G.42d, with that obtained by R. Jones-Phillipson for Zulu and Xhosa, S.42 and S.41. (Op.cit.) They are 2089 and 2310 respectively. While the former pair are dialects of the same language, the latter have always been considered as two distinct languages.



Proportional illustration of Common Bantu reflexes in each pair of Selected Languages.

(Prepared by M.Mann from figures supplied by M.Slavíková.)

Overlapping areas of each two diamonds indicate the number of reflexes occurring jointly in two languages; excess area of right-hand diamond indicates the number of reflexes unique to language on the right; excess area of left-hand diamond indicates the number of reflexes unique to language on the left.

Another example of a relatively high order of similarity as suggested by a large number of shared indexes, and countered by relatively large numbers of unique reflexes is that of Kikuyu and Unguja. Theirs is the fourth highest number of shared reflexes. With the overall highest total number of unique reflexes, however, they are fifteenth in order of similarity; 205 unique to Kikuyu and 458 to Unguja make 663.

In the table below are presented the figures relevant to the arguments presented above.

Languages investigated X+Y	Number of shared reflexes	Nu. of refl. peculiar to X	Nu. of refl. peculiar to Y	IR
M+U	507	56	158	2089
K+U	205	205	458	198
K+M	170	228	390	170
S+K	130	168	269	185

The last example in the table, that of Saghala and Kikuyu, demonstrates how the lowest obtained number of shared reflexes need not indicate the least degree of relationship. In spite of the mere 130 reflexes they have in common Saghala and Kikuyu do not reach a degree of separate uniqueness (as do for example Kikuyu and Unguja) which would reduce the IR figure more dramatically and make it correspond simply with the number of shared reflexes.

Results of the comparisons between the IR figures and those expressing the numbers of shared reflexes are laid out in the two following tables. All Selected Languages had been compared among themselves and ordered according to the two indices.

Both tables contain an extra column headed SI (Simple Index). It is a simplified variety of Guthrie's IR, and will be discussed in the section immediately following the tables.

Selected Languages ordered according to
1/ IR, 2/ number of jointly occurring reflexes,
3/ SI.

	Languages investigated	IR	Languages investigated	Joint reflexes	Languages investigated	SI
1	M+U	2089	M+U	507	M+U	703
2	G+M	737	G+U	348	G+M	493
3	D+S	642	G+M	333	D+S	466
4	G+U	612	K+U	205	G+U	443
5	S+G	364	S+G	196	S+G	341
6	D+G	294	D+S	188	D+G	301
7	S+M	247	S+U	181	S+M	265
8	D+K	221	S+M	180	D+K	256
9	G+K	211	D+G	172	G+K	251
10	S+U	207	K+M	170	K+U	236
11	K+U	198	G+K	168	S+U	231
12	S+K	185	D+U	165	S+K	229
13	D+M	177	D+M	151	K+M	216
14	D+U	171	D+K	141	D+M	215
15	K+M	170	S+K	130	D+U	205

Statistical classification presented above is summarized in the next table, demonstrating how different attributes may affect the hierarchy of relationships.

Place of order according to the IR, number of jointly occurring reflexes, and SI.

	Languages	IR	Joint reflexes	SI
1	M+U	1	1	1
2	G+M	2	3	2
3	D+S	3	6	3
4	G+U	4	2	4
5	S+G	5	5	5
6	D+G	6	9	6
7	S+M	7	8	7
8	D+K	8	14	8
9	G+K	9	11	9
10	S+U	10	7	11
11	K+U	11	4	10
12	S+K	12	15	12
13	D+M	13	13	14
14	D+U	14	12	15
15	K+M	15	10	13

b) Simple index

This is a simple index of relationship obtained through the formula $q/(x+y+q)$, where, as in the IR formula, 'q' represents the number of shared reflexes, 'x' the number of reflexes exclusive to the first language, and 'y' the number of reflexes exclusive to the second language of the pair in question. Index figures obtained for each pair of the Selected Languages are set out in the SI column in the above tables.[1]

---

[1] For the sake of convenience the final figures were multiplied by 1000.

When figures from the following table were plotted on a map (1 = 1cm) with languages located relative to their present positions, the IR figures yielded a picture which corresponded more nearly to present topography than did the SI figures.

Relative degrees of relationship between pairs of Selected Languages, expressed in multiples of Mvita+Unguja relationship.

Order	Languages investigated	IR	SI
1	M+U	1	1
2	G+M	2,8	1,4
3	D+S	3,2	1,5
4	G+U	3,4	1,6
5	S+G	5,7	2,1
6	D+G	7,1	2,3
7	S+M	8,4	2,6
8	D+K	9,4	2,7
9	G+K	9,9	2,8
10	S+U	10,1	3,0
11	K+U	10,5	2,97
12	S+K	11,3	3,1
13	D+M	11,8	3,3
14	D+U	12,2	3,4
15	K+M	12,3	3,3

Note: In using qualifying expressions which describe relationships between languages as being 'relatively' or 'comparatively' 'close' or 'remote', references are implied to the remaining Selected Languages only.

3.1.3 A note on skewed reflex statistics

Parallel to direct reflexes three relationship categories were established from the skewed reflex data: reflexes shared by the two languages in question, those exclusive to the first language, and those exclusive to the second language.

However, it is their qualitative heterogeneity that demands a special treatment of skewed data as markers of relationship.

While two direct reflexes of a starred form are members of a narrowly defined category, their skewed counterparts may, but need not be (as is the majority) of a certain type - therefore not always comparable on a single level.

It would therefore be inadequate for direct and unqualified inferences to be made on the statistical evidence alone of skewed reflexes, such as is presented in this work. (See the general statistical table of all direct and skewed cognates, p.51.) A massive commentary would be required if conclusive inferences were to be drawn from skewed cognate statistics.

Let us consider a straightforward synthesis of direct and skewed data, both of which happen to produce an identical type of results.

On all available evidence Mvita+Unguja display the closest relationship among the Selected Languages.

1) Their IR and SI are by far the highest.
2) They share the largest number of direct reflexes.
3) Mvita has the overall lowest number of exclusive reflexes while Unguja's figure is among the six lowest out of a total of fifteen.
4) They share by far the largest number of skewed reflexes.

5) The total number of skewed reflexes exclusive to each language is the lowest of all pairs. On the evidence of the SI formula applied to the skewed reflexes the result was by far the highest figure, 703, the immediate lower figure being 493.

Nevertheless, even such unequivocal evidence as that of Mvita+Unguja is not free from the intrinsic element of misrepresentation of its skewed data. For the figures in 4) and 5) represent skewings not only of different types and degrees but of essentially different origins or geographical spread.

3.2 COMMON BANTU REFLEXES SHARED BY DAWIDA AND OTHER SELECTED LANGUAGES

The first five relationships to be discussed will be as follows,

1 Dawida and Saghala

2 Dawida and Giryama

3 Dawida and Kikuyu

4 Dawida and Mvita

5 Dawida and Unguja,

as demonstrated in the 'Triangle of Relationships' at the beginning of this chapter. All the remaining Selected Languages and their relationships will be presented in identical fashion in the following section.

Moreover, the same order of presentation will be adopted throughout this chapter, and will therefore be quoted in full only in this introductory section.

Order of presentation

3.2.1 Statistical data presented

3.2.1.1 Statistical data for the IR

3.2.2 Individual statistical characteristics

3.2.2.1 Shared reflexes

3.2.2.2 Unique reflexes

3.2.3 Summary of statistical data interpreted

3.2.4 Inferences from statistical characteristics

3.2.5 Conclusions

3.2.1 Statistical data presented

3.2.1.1 Statistical data for the IR

In the table below is presented the hierarchy[1] of closeness of relationship between Dawida and the remaining Selected Languages, established on the basis of the Index of Relationship (IR) figures, together with the order of difference.[2]

	Languages	IR	% difference from preceding pair
1	Dawida+Saghala	642	-
2	Dawida+Giryama	294	54.2
3	Dawida+Kikuyu	221	24.8
4	Dawida+Mvita	177	19.9
5	Dawida+Unguja	171	3.4

The order of difference in the IR among pairs of Selected Languages involving Dawida.

3.2.2 Individual Statistical Characteristics

We shall now consider to what extent the two basic components of the IR, reflexes shared and reflexes unique, correspond to or differ from the conditions expressed by the IR figures.

---

[1] 'Hierarchy' used throughout this work where 'rank order' might be used by others.

[2] The 'steps' between successive IR figures are given in % in order to express geometrical rather than arithmetic progression.

3.2.2.1 Shared reflexes

	Languages	Shared reflexes	Difference from preceding pair
1	Dawida+Saghala	188	-
2	Dawida+Giryama	172	16
3	Dawida+Unguja	165	7
4	Dawida+Mvita	151	14
5	Dawida+Kikuyu	141	10

Table Degrees of difference in the number of shared reflexes among pairs of Selected Languages involving Dawida.

The differences among pairs of languages are more or less of the same order, with the exception of Dawida+Unguja. According to this index Dawida and Unguja have a relationship not too dissimilar from that of Dawida and Giryama.

This state of affairs is, however, at some variance with the IR results.

3.2.2.2 Unique reflexes

	Languages	Unique to 1st	Unique to 2nd	Total	Difference from preceding pair
1	Dawida+Saghala	105	110	215	-
2	Dawida+Giryama	116	283	399	184
3	Dawida+Kikuyu	146	263	409	10
4	Dawida+Mvita	142	408	550	141
5	Dawida+Unguja	133	506	639	89

Table Degrees of difference in the number of unique reflexes among pairs of Selected Languages involving Dawida.

By far the greatest numerical distance exists between the first two pairs, Dawida+Saghala and Dawida+Giryama.

The second largest gap lies between Dawida+Kikuyu and Dawida+Mvita, i.e. the third and fourth pairs.

A very small difference exists between the second and third pairs, Dawida+Giryama and Dawida+Kikuyu, but the fact that there is a much larger difference between the figures for unique reflexes for the former pair of reflexes, is significant. On the basis of these differences between the numbers of languages can be assorted into three broad groupings separated from each other by substantial numerical gaps:

1. Dawida and Saghala
2. Giryama and Kikuyu
3. Mvita and Unguja

The number of reflexes unique to Dawida in relation to the other languages increases regularly and in the same order as its shared reflexes. The order is as follows: 1) Saghala, 2) Giryama, 3) Unguja, 4) Mvita, 5) Kikuyu. Their opposite counterparts, however, reverse this into the resulting order presented in the above table.

3.2.3 Summary of statistical results interpreted

The results of the statistical investigations can be summarized as follows:

a) Dawida and Saghala have the highest, which is also 'simple'.[1] They have the highest IR figure, the largest number of direct reflexes in common, and the least number of direct reflexes unique to each.

---

[1] We shall term 'simple' such IR where all of its constituents are of the same hierarchical order relative to other Selected Languages.

All their statistical characteristics are expressed in figures substantially higher (or lower) than those of the remaining pairs. This is particularly so in the case of the IR, 642, with the immediate lower figure 294 (54.2% difference).

b) Dawida and Unguja have the lowest IR at 171. It is, in fact, the second overall lowest IR obtained between two Selected Languages, and the 506 reflexes unique to Unguja, a strong influencing factor, represent the highest figure obtained for any language in this study.

c) With respect to the IR results, there is a reversal of order between Unguja and Kikuyu in the shared reflex hierarchy in which, with an index figure of 165, Unguja comes uppermost.

d) The IR's of the first two pairs are widely different. The IR of Dawida and Giryama is appreciably lower than that of Dawida and Saghala, 294 and 642 respectively.

There is also a sizable difference of 24.8% between the second and third pairs, the two indices being 294 and 221 for Dawida+Giryama and Dawida+Kikuyu respectively, yet it is much smaller than the gap separating the first two pairs (54.2%).

The difference between the last and the penultimate pairs is only 3.4%: Dawida+Mvita's IR is 177, Dawida+Unguja's 171.

e) The differences among pairs in their IR figures can be used as a basis for grouping the Selected Languages.

Thus Dawida and Saghala would form the <u>first group</u>, with an IR substantially higher than any other pair. Giryama would be the sole member of the <u>second group</u>, having an IR much lower than the first pair - but much higher on the other hand than the remaining pairs. The <u>third</u> broad <u>grouping</u> would consist of

Kikuyu, Mvita and Unguja, all having appreciably lower IR's with Dawida than the previous pairs, but with differences between them smaller than those which separate the individual groups just outlined, especially in the case of Unguja (3.4%).

3.2.4 Inferences from statistical characteristics

Several principle inferences were drawn from the selected statistical characteristics of the data. Its most explicit part referred to

1) Dawida and Saghala. It appears that of the five investigated relationships of Dawida with the other Selected Languages, theirs is by far the closest.

On all available evidence, and this includes even the evidence of skewed cognates, the indices obtained indicate a straightforward relationship with an affinity far closer than between the four other pairs.

Their IR of 642 is high enough for an immediate link in prehistory between them to be postulated.

2) Dawida and Giryama. Their relationship is substantially remoter with an IR of 294. This considerable drop is due to the large number of reflexes exclusive to Giryama. While Dawida has a number of unique reflexes which is not too dissimilar from that it has with respect to Saghala, Giryama has a very large number of reflexes unique to itself. Notwithstanding the fact that these first two pairs, Dawida+Saghala and Dawida+Giryama, have similar numbers of shared reflexes (188 and 172), the weight of its unique reflexes places Giryama outside the zone of an immediate link with Dawida in prehistory.

3) Dawida and Kikuyu, with an IR of 221, display less similarity still than the two previous pairs. It is interesting to note that they share the least number of reflexes, but on the other hand their unique reflexes are not particularly numerous: a statement valid especially for Kikuyu.

Their comparatively low degree of affinity makes it improbable for Dawida and Kikuyu to have shared a common immediate ancestor at a stage in prehistory when Dawida and Saghala might have done so.

4) and 5) Dawida and Mvita, and Dawida and Unguja. The two relationships are marked by the largest numbers among all five pairs of unique reflexes. In both cases it is the Swahili element that is responsible, indicating that some focus of its relationships lies outside the zone of Dawida. Of the two dialects Unguja has the significantly higher figure, thus being placed as the last in order of closeness of relationship with Dawida. The numerical difference between the pairs is small: their IR's are 177 and 171 respectively,[1] a difference of 3.4%. An immediate link in prehistory with Dawida cannot be postulated for either Swahili dialect.[2]

3.2.5 Conclusions

The data bear out the present relationship of Dawida and Saghala as an essentially close one, with a possible immediate link in prehistory.

---

[1] For further discussion of the Dawida-Unguja relationship see Section 3.7 below.

[2] Implying that other, intermediate, links may have existed.

Theirs is the unambiguously closest relationship among all the five pairs centred round Dawida, thus separating them from the remaining languages.

The next closest relationship is that of Dawida and Giryama. It is, however, much lower in rank, and does not indicate a possibility of an immediate link in prehistory.

The third pair in order of closeness of relationship, Dawida and Kikuyu, have a remoter relationship still - and suggest by this result that Dawida ought not to be classified as an immediate part of the 'up-country' languages.

The last two pairs, however, indicate that neither ought it to be included among the 'coastal' languages, its higher IR with Giryama notwithstanding. The rather remote relationship of Dawida with the two Swahili dialects indicates the improbability of an immediate link in prehistory between them.

3.3 COMMON BANTU REFLEXES SHARED BY SAGHALA AND OTHER SELECTED LANGUAGES

3.3.1 Statistical data presented

3.3.1.1 Statistical data for the IR

The following table demonstrates the hierarchy among the pairs of Selected Languages involving Saghala, established on the basis of the IR figures. The last column indicates the differences in %.

	Languages	IR	% difference from preceding pair
1	Saghala+Dawida	642	-
2	Saghala+Giryama	364	43.3
3	Saghala+Mvita	247	32.1
4	Saghala+Unguja	207	16.2
5	Saghala+Kikuyu	185	10.6

Table Hierarchy of closeness of relationship between Saghala and other Selected Languages, according to the IR.

3.3.2 Individual statistical characteristics

3.3.2.1 Shared reflexes

	Languages	Shared reflexes	Difference from preceding pair
1	Saghala+Giryama	196	-
2	Saghala+Dawida	188	8
3	Saghala+Unguja	181	7
4	Saghala+Mvita	180	1
5	Saghala+Kikuyu	130	50

Table The order of difference in the number of shared reflexes among pairs of Selected Languages involving Saghala.

3.3.2.2 Unique reflexes

	Languages	Unique to 1st	Unique to 2nd	Total unique	Difference from preceding pair
1	Saghala+Dawida	110	105	215	-
2	Saghala+Giryama	109	270	379	164
3	Saghala+Kikuyu	168	269	437	58
4	Saghala+Mvita	115	385	500	63
5	Saghala+Unguja	114	487	601	101

Table The order of difference in the number of unique reflexes among pairs of Selected Languages involving Saghala.

3.3.3 Summary of statistical results interpreted

These are the principal results of the statistical investigation of the pairs of Selected Languages focussed on Saghala:

a) Saghala and Dawida have the highest IR, but not a 'simple' one: Saghala shares the highest number of direct reflexes not with Dawida but with Giryama. The difference between the two is, however, only eight points.

b) The lowest IR was obtained for Saghala and Kikuyu: 185, although they do not have the largest total number of unique reflexes.

c) The IR decreases fairly rapidly down the hierarchy. The degrees of difference between pairs are substantial but none reaches the level of difference between Dawida+Saghala and Dawida+Giryama, 54.2%.

- By far the greatest difference exists between the first two pairs, Saghala+Dawida and Saghala+Giryama, as demonstrated by the IR, the difference in 43.3%, and by the unique index figures.

- There is also clear difference between the two pairs involving

Swahili, with the exception of the shared reflex figures. In each case it is Mvita which scores the higher figure (16.2% difference in the IR).

d) The IR of the last and the penultimate pairs are not very widely different. Saghala+Unguja and Saghala+Kikuyu display 10.6% difference between their IR's.

3.3.4 Inferences from statistical characteristics

A number of basic inferences were drawn from the statistical data. They can be summed up as follows:

1) The closest relationship can be inferred for Saghala and Dawida. They have the highest IR, 642. However, this relationship is not identical with the Dawida-Saghala one in spite of the same IR. Saghala proved to be closer to Giryama in one aspect, the shared reflex index.

2) Saghala and Giryama display the second closest relationship, although with an IR much lower than the preceding pair's. The gap between them of 43.3% is due to the large number of reflexes unique to Giryama (270) as opposed to 105 only that are unique to Dawida.

It suggests that the generally assumed 'inclination' of Saghala towards Giryama and other Miji-kenda languages be viewed with some circumspection and, furthermore, that the inferred relationship is not close enough for an immediate link between the two languages to have existed in prehistory.

3) and 4) Saghala+Mvita and Saghala+Unguja have the next remoter relationships, as expressed by their IR's of 247 and 207 respectively. Both are well below the two preceding IR's, and contain

in them the largest numbers of unique reflexes.

The closer relationship with Saghala together with the extra proximity factor[1] place Mvita before Unguja, there being an appreciable difference of 16.2% between them.

Both relationships are indicative of separate developments of the languages they contain within the late period of PB-B.[2]

5) Saghala and Kikuyu are last in the hierarchy, with an IR of 185. It indicates a basically remote relationship within the context of the Selected Languages. Kikuyu stands out in two cases both of which suggest its isolation from the remaining languages.

Firstly, it shows the only outstanding dissimilarity in the small number of reflexes it shares with Saghala. Secondly, Saghala has the largest number of reflexes unique to itself in relation to Kikuyu. The figure, 168, is *much higher than* the rest, which range between 109 and 115.

In conclusion, no immediate link in prehistory between the two languages can be inferred but a fission that may have occurred prior to separation of Saghala with all the other Selected Languages.

3.3.5 Conclusions

Similarly to the previous Section on Dawida and its relationships with other Selected Languages, the closest relationship indicated by

---

[1] For discussion of the 'proximity factor' see CompB, Vol.1, 63.73/a/. It expresses the phenomenon of the IR decreasing with distance.

[2] For discussion of the PB-B period, the era after Proto-Bantu had split into two dialects of which the PB-B denotes the eastern one, see CompB, Vol.2, Chap.8 and Appendix 8/3, 'Framework of a possible Bantu genealogy.'

the data in this Section is that between Saghala and Dawida. Both their IR (642) and the added proximity factor make it feasible for an immediate link in prehistory between them to be postulated. Their fission may be inferred as occurring at a relatively 'recent' stage of the PB-B period.

A slight, if not negligible linkage between the first two pairs of the hierarchy of closeness of relationship is indicated by the largest number of reflexes that is shared between Saghala and Giryama. Nevertheless, although their relationship is not a remote one, an immediate link in prehistory would be a matter of some circumspect chronology.

The two Swahili dialects are separated in their not particularly close relationships with Saghala by a large difference of the individual numbers of their unique reflexes, with Mvita displaying the closer relationship with Saghala. However, neither of the Swahili dialects appear to be likely to have had an immediate link in prehistory with Saghala.

Kikuyu appears to be isolated from the rest on most evidence. It has the lowest IR with Saghala among the five pairs, 185, and the individual evidence of reflexes shared and unique also excludes a probability of an immediate connection with Saghala in prehistory. According to the results of the present Section, the first fission to have taken place between Saghala and another Selected Language from a common ancestor in an early period of PB-B can be postulated for Kikuyu.

To conclude, Saghala appears, on evidence of the present Section, to be more 'coastal' than Dawida. It has a closer relationship with Giryama than Dawida does (19.2% difference), a closer relationship still with Mvita (28.3% difference), and with Unguja (17.4% difference).

On the other hand, Saghala has a remoter relationship with Kikuyu than Dawida (16.3% difference).

3.4 COMMON BANTU REFLEXES SHARED BY GIRYAMA AND OTHER SELECTED LANGUAGES

3.4.1 Statistical data presented

3.4.1.1 Statistical data for the IR

	Languages	IR	% difference from preceding pair
1	Giryama+Mvita	737	-
2	Giryama+Unguja	612	17
3	Giryama+Saghala	364	40.5
4	Giryama+Dawida	294	19.2
5	Giryama+Kikuyu	211	28.2

Table : Hierarchy of closeness of relationship according to the IR.

3.4.2 Individual statistical characteristics

3.4.2.1 Shared reflexes

	Languages	Shared reflexes	Difference from preceding pair
1	Giryama+Unguja	348	-
2	Giryama+Mvita	333	15
3	Giryama+Saghala	196	137
4	Giryama+Dawida	172	24
5	Giryama+Kikuyu	168	4

Table : Degrees of difference in the number of shared reflexes among pairs of Selected Languages involving Giryama.

3.4.2.2 Unique reflexes

	Languages	Unique to 1st	Unique to 2nd	Total unique	Difference from preceding pair
1	Giryama+Mvita	128	214	342.	-
2	Giryama+Saghala	270	109	379	37
3	Giryama+Dawida	283	116	399	20
4	Giryama+Unguja	112	324	436	37
5	Giryama+Kikuyu	274	225	499	63

Table : Degrees of difference in the number of unique reflexes among pairs of Selected Languages involving Giryama.

3.4.3 Summary of statistical results interpreted

The data contained in the above tables can be summarized in the following points:

a) Although Giryama and Mvita have the highest IR, 737, their relationship is not formed by characteristics of the same hierarchical order. With respect to the IR, shared reflex index, and unique reflex index, Mvita's position is slightly different in each case, yet the IR figure is relatively high.

b) Giryama and Kikuyu have the lowest, and 'simple', IR at 211. All their indices are expressed in the smallest figures - or highest in the case of unique reflexes.

c) The differences in the IR between pairs are rather uneven, except in the shared reflex hierarchy (with Kikuyu 25 an 'irregularity'). An obvious dividing line is to be drawn between the two Swahili dialects and Giryama on one side and the remaining languages on the other. The dividing line is provided by a large gap in the IR values - 40.5% difference.

Kikuyu, the last in the hierarchy of closeness, is separated from the rest by one measure in particular: the very few reflexes it shares with Giryama.

d) By far the highest IR's are reached between Giryama and the two Swahili dialects. Mvita and Giryama have in fact the overall highest IR discussed so far, 737. On all evidence, except Unguja's unique reflexes, the Swahili dialects take up the first two positions in the hierarchy of closeness of relationship, and remain separated from the other languages by a large difference in the IR.

3.4.4 Inferences from statistical characteristics

Several principal inferences were drawn from the results of the statistical data, which can be summarized in the following points:

1) Among the five pairs of languages investigated in this Section Giryama and Mvita display the closest relationship. Theirs is the closest relationship propounded in this work so far for two Selected Languages. Their IR is 737. The two immediate lower figures, 642 and 612, were obtained between Dawida+Saghala and Giryama+Unguja respectively. The high IR is suggestive of an immediate link in prehistory. Notwithstanding this hypothesis, however, the relatively substantial number of reflexes unique to Mvita (214) in contrast to Giryama's 128, does indicate that Mvita has an alternative focus of its relationship outside the affinity with Giryama.

2) Giryama and Unguja appear to have also a close relationship with each other, yet not at all as close as the preceding pair. (There is 17% difference in the IR between the pairs.)

Unguja, like Mvita, scored a very high number of unique reflexes in relation to Giryama's few, thus indicating a focus of Unguja's relationships elsewhere.

Nevertheless, the IR figure as well as the shared reflex index are both high enough for an immediate link in prehistory to be postulated between the two languages.

3) Giryama and Saghala are third in the hierarchy, separated from the preceding two pairs by a percentual distance of 40.5, which indicates a rift in relationships among the Selected Languages.

Although Giryama+Saghala have a relatively close relationship among the Selected Languages, their IR is 364, it would be highly speculative to propound an immediate link between them in prehistory.

4) Giryama and Dawida, with an IR of 294, display not a particularly close relationship. Their IR is by 19.2% lower than the immediately preceding one. The centrifugal element here is Giryama with its large number of unique reflexes, which suggest closer relationships away from Dawida.

The character of the Giryama+Dawida relationship makes it unlikely for an immediate link in prehistory to have existed between them.

5) Giryama and Kikuyu appear to have a rather remote relationship, without an immediate connection in prehistory. Their IR is not high at 211, and both languages tend towards opposite directions in their affinites, as inferred from the equally large numbers of their respective unique reflexes.

6) Although the 'rift' separating Giryama and Swahili from the rest is appreciable (40.5% difference), it is less than the 54.2% and 43.3% separating Dawida and Saghala from the remaining languages.

3.4.5 Conclusions

The results concerning closeness of the five relationships between Giryama and the remaining Selected Languages yield such conclusions as indicate a dividing line between Giryama and the two Swahili dialects on one side, and Saghala, Dawida and Kikuyu with not so close relationships with Giryama on the other.

Of the two Swahili dialects, Mvita has a much closer relationship with Giryama than does Unguja. Their immediate links in prehistory with Giryama are likely to have been severed at two different points on the relative time-scale. Their individual unique-reflex indices suggest a tendency of Mvita and Unguja towards alternative foci of relationship away from Giryama.

Giryama's relationship with Saghala is not a particularly close one, but neither is it remote. To speculate about their link in prehistory would necessitate further comparative data together with some amount of precarious chronology. At any rate, though, Giryama's relationship with Saghala is closer than with Dawida.

Dawida appears to be more closely related to Saghala and Kikuyu - a state of affairs that lends support to the diachronic hypothesis of a dividing line between the two tentative groupings outlined above.

Giryama then does not enjoy a close relationship with Dawida, and even less so with Kikuyu. The latter's affinity with Giryama is expressed by the lowest 'simple' IR among the five pairs: all of its constituents are expressed in figures that indicate the remotest relationship.

Neither Dawida nor Kikuyu are likely to have had an immediate connection with Giryama in prehistory.

3.5 COMMON BANTU REFLEXES SHARED BY KIKUYU AND OTHER SELECTED LANGUAGES

3.5.1 Statistical data presented

3.5.1.1 Statistical data for the IR

As the first step in the discussion of relationships of the Selected Languages with Kikuyu at its centre, we shall present the direct reflex data relevant to the IR.

	Languages	IR	% difference from preceding pair
1	Kikuyu+Dawida	221	-
2	Kikuyu+Giryama	211	4.5
3	Kikuyu-Unguja	198	6.2
4	Kikuyu+Saghala	185	6.6
5	Kikuyu+Mvita	170	8.1

Table : Hierarchy of closeness of relationship between Kikuyu and other Selected Languages according to the IR.

3.5.2 Individual statistical characteristics

3.5.2.1 Shared reflexes

	Languages	Shared reflexes	Difference from preceding pair
1	Kikuyu+Unguja	205	-
2	Kikuyu+Mvita	170	35
3	Kikuyu+Giryama	168	2
4	Kikuyu+Dawida	141	27
5	Kikuyu+Saghala	130	11

Table : Degrees of difference in the number of shared reflexes among pairs of Selected Languages involving Kikuyu.

3.5.2.2 Unique reflexes

	Languages	Unique to 1st	Unique to 2nd	Total unique	Difference from preceding pair
1	Kikuyu+Dawida	263	146	409	-
2	Kikuyu+Saghala	269	168	437	28
3	Kikuyu+Giryama	225	274	499	62
4	Kikuyu+Mvita	228	390	618	119
5	Kikuyu-Unguja	205	458	663	45

Table : Degrees of difference in the number of unique reflexes among pairs of Selected Languages involving Kikuyu.

3.5.3 Summary of statistical results interpreted

Since the IR figures discussed below are essentially low, terms like 'high', 'great', and other similar ones refer to them in relation to the other pairs of Selected Languages of this Section only.

a) The highest IR, 221, was obtained between Kikuyu and Dawida. They share a small number of reflexes, 141, but their unique index reflex was decidedly the lowest at 409.

b) The lowest IR was obtained for Kikuyu and Mvita, 170, which is the overall lowest figure obtained among the Selected Languages.

Mvita has a very large number of unique reflexes (390) but is second in order of shared reflex numbers with 170 points.

c) The differences between pairs according to their IR's are remarkably even - 4.5 to 8.1% -, relatively small, and without a single irregularity.

The substantial gap existing in the unique reflex hierarchy was neutralized within the IR formula.

d) The position of Saghala is surprisingly low in the hierarchy of closeness of relationship, with an IR of 185.

Unguja's position as third in order is higher than expected, with an IR of 198.

e) All five IR's investigated in this Section are very low. Indeed, Giryama, Saghala and Mvita have their lowest IR's with Kikuyu of all Selected Languages, thus forming a distinct group.

3.5.4 Inferences from statistical characteristics

The following points will summarize the principal inferences concerning the Selected Languages focussed on Kikuyu that were drawn from selected statistical data.

1) All the five relationships centred round Kikuyu are rather remote. Indeed, this Section includes the remotest relationships among the Selected Languages.

Hence the inference concerning the isolation of Kikuyu from the remaining languages and made in previous Sections, is supported by results of the present Section.

2) On account of the essentially low IR figures, no immediate link in prehistory can be postulated between any two pairs.

3) The closest among the relatively remote relationships is that of Kikuyu and Dawida, with an IR of 221. Their geographical proximity adds extra strength to this inference.

4) Kikuyu and Giryama have the next remoter relationship. Giryama appears closer to Kikuyu than the other languages in the two aspects of shared and unique reflexes.

It is of some interest that a closer affinity should be postulated between Giryama and Kikuyu (IR = 211) than between Saghala

and Kikuyu (IR = 185) since this result conflicts with the 'proximity factor'.

5) Kikuyu and Unguja, with an IR of 198, rank among the remote relationships between the Selected Languages, although they share a substantial number of direct reflexes (205).

Next after Unguja in order of closeness is not Mvita, in consistency with the hierarchies in the previous Sections, but

6) (Kikuyu and) Saghala (IR = 185). It is worth noting that the position of Saghala is interspersed between the two Swahili dialects, and may indicate that a split of Kikuyu and Saghala occurred after its split with Mvita but before the split with Unguja on the relative time-scale.

7) Kikuyu and Mvita (IR = 170) is the remotest of all relationships among the Selected Languages, but not a 'simple' one.

3.5.5 Conclusions

The Selected Languages focussed in their relationships on Kikuyu all yielded low IR figures, thus indicating remote relationships with it. Indeed, Kikuyu's relationships with the other languages are characterized by the lowest IR's obtained for a Selected Language.

As shown in previous Sections, an isolation of Kikuyu from the other languages is feasible to postulate from the statistical data. The low IR figures suggest that a split probably took place at a relatively early stage of the PB-E period, which separated Kikuyu from the group formed by the remaining Selected Languages, with Dawida separating 'soon' after according to the relative time-scale.

Differences in closeness between individual pairs are small and identical in rank, as indicated by the differences in the IR of between

10 and 15 IR points only.

Kikuyu displays closest affinity to Dawida, IR = 221, followed by Giryama, IR = 211. By implication, Kikuyu appears to be more closely related to Giryama than to Saghala, the substantially close relationship of Saghala and Dawida notwithstanding.

Third in order of closeness is the relationship with Unguja, IR = 198. It is also a substantially remote relationship, but a closer link is thus indicated between Kikuyu and Unguja than between Kikuyu and Saghala. Unguja's position in the middle of the hierarchy rather than at its end is contrary to the condition described as the 'proximity factor'. Theirs is the second relationship in this group that counters the generally valid condition of the IR decreasing with distance.

Hence Mvita is found relegated to the last position in the hierarchy, having the remotest relationship with Kikuyu and the overall lowest IR of 170.

No groupings of the Selected Languages can be established on the basis of differing degrees of closeness as the differences encountered in this Section are negligible for this purpose.

3.6 COMMON BANTU REFLEXES SHARED BY MVITA AND OTHER SELECTED LANGUAGES

3.6.1 Statistical data presented

3.6.1.1 Statistical data for the IR

	Languages	IR	% difference from preceding pair
1	Mvita+Unguja	2089	-
2	Mvita+Giryama	737	64.7
3	Mvita+Saghala	247	66.5
4	Mvita+Dawida	177	28.3
5	Mvita+Kikuyu	170	4

Table : The hierarchy of closeness of relationship between Mvita and the other Selected Languages according to the IR.

3.6.2 Individual statistical characteristics

3.6.2.1 Shared reflexes

	Languages	Shared reflexes	Difference from preceding pair
1	Mvita+Unguja	507	-
2	Mvita+Giryama	333	174
3	Mvita+Saghala	180	153
4	Mvita+Kikuyu	170	10
5	Mvita+Dawida	151	19

Table : Degrees of difference in the number of shared reflexes among pairs of Selected Languages involving Mvita.

3.6.2.2 Unique reflexes

	Languages	Unique to 1st	Unique to 2nd	Total unique	Difference from preceding pair
1	Mvita+Unguja	56	158	214	-
2	Mvita+Giryama	214	128	342	128
3	Mvita+Saghala	385	115	500	158
4	Mvita+Dawida	408	142	550	50
5	Mvita+Kikuyu	390	228	618	68

Table : Degrees of difference in the number of unique reflexes among pairs of Selected Languages involving Mvita.

3.6.3 Summary of statistical results interpreted

With the exception of Kikuyu and Dawida the order of individual pairs is identical in all three tables that were presented above. Their IR's can therefore be termed 'simple'.

Interpretation of the statistical results is summarized in the following points:

a) Mvita and Unguja have by far the highest IR at 2089. It is not only the overall highest IR among the Selected Languages, but it is also unique in being so high. The immediate lower figure is 737.

b) The lowest IR is that of Mvita and Kikuyu, 170, which is the overall lowest figure obtained in this study. The reflexes they share are not particularly numerous while their respective unique reflexes are extremely so.

c) As stated in the discussion of shared and unique reflex hierarchies, both their order and their internal composition are nearly identical. Only the last two pairs, Mvita+Dawida and Mvita+Kikuyu have results that are contrary to this condition.

The IR results broadly correspond with this situation.

d) There are two ranks of numerical distances in the IR as well as in the shared and unique reflex tables: 'large', and 'relatively small'. The 'large' IR gaps are much larger than those of its two constituent tables.

e) The two ranks of figures indicate a division into groups of 1. high, and 2. low IR's. The groups consist of : 1. Mvita+Unguja, Mvita+Giryama, and 2. Mvita+Saghala, Mvita+Dawida and Mvita+Kikuyu.

The group having the high IR's could be divided further due to the substantial difference between the two pairs of 64.7%.

3.6.4 Inferences from statistical characteristics

The following are the principal inferences drawn from the statistical data on the Selected Languages centred round Mvita:

1) With the exception of Dawida and Kikuyu 'simple' IR's were obtained between Mvita and the other Selected Languages, which indicates unambiguous and straightforward relationships.

2) Mvita and Unguja display the evidently closest relationship inferred so far. Their IR of 2089 outnumbers by almost three times the immediate lower one.

All markers of relationship point to such a high measure of closeness, which in turn can be interpreted as indicating a strong likelihood of an immediate link in prehistory between them.

Furthermore, one can infer from the exceptionally high degree of closeness that the split between Mvita and Unguja occurred rather recently on the relative time-scale.

3) Mvita and Giryama's IR is 737, and although it represents a relatively very close relationship, it is nowhere near that of Mvita and Unguja.

However, this, too, is a 'simple' affinity in which all characteristics point to the same position in the hierarchy, in this case the second in order and substantially close in character. An immediate link in prehistory between Mvita and Giryama can easily be inferred from the evidence.

4) Mvita and Saghala's relationship seems a degree less close than that of Mvita and Giryama. The IR figure is not particularly high at 247, and Mvita displays a substantial degree of exclusiveness with 385 reflexes unique to itself.

Saghala's relationship *to* Mvita may be described as a not particularly close one, without a probability of an immediate link in prehistory.

5) Mvita and Dawida can be said to have a somewhat remote affinity within the context of Selected Languages. Their IR is appreciably low at 177 in spite of the fact that Dawida seems to be closely related to Saghala, which in turn displays a closer affinity with Mvita.

No immediate link in prehistory is postulated on the evidence investigated.

6) Mvita and Kikuyu 'are' the least related pair in this Section and, indeed, in the whole study. Theirs is the overall lowest IR at 170, which lends itself to no postulation of an immediate link in prehistory. It is, however, not a 'simple' remote relationship, viewed from either of the two languages.

7) There are substantial differences in the closeness of the discussed relationships, which indicate a division of the languages into the following groups:

A Mvita and Unguja form a group of their own, having a much closer affinity that the rest;

B Giryama, being separated from the preceding two by a large difference, yet still rather close in relationship to both of them;

C Saghala, Dawida, Kikuyu. The three are separated from the preceding by a sizable difference, while differences between them are comparatively small.

3.6.5 Conclusions

The statistical evidence of the five relationships centered round Mvita reveals relationships that differ very widely in their degrees of closeness. This, in turn, implies very widely separated points on the relative time-scale at which the languages may have split.

The data bear out two broad groupings.

Languages in the first group may be said to be closely related to each other, are likely to have had immediate links in prehistory with each other, and 'rather recent' fissions from their common immediate ancestor in the PB-B or the Bantu period. Mvita, Unguja and Giryama belong to this group although Giryama, on stricter criteria, might be separated on account of its relationship with Mvita being substantially less close.

The second group includes Saghala, Dawida, and Kikuyu. Neither has a close relationship with Mvita, and each in turn has a remoter relationship with it than the preceding one, Mvita and Kikuyu having the remotest relationship among the

Selected Languages. No immediate link in prehistory is therefore postulated between any of them and Mvita, but a separation from their common intermediate ancestor, which occurred at a 'much earlier' date in PB-B than the separation of the first group.

On a more rigorous differentiation the second group would have to be divided in to Saghala as a single-member group on one side, with a clearly closer affinity with Mvita, and Dawida and Kikuyu on the other with a remoter relationship with Mvita.

3.7 COMMON BANTU REFLEXES SHARED BY UNGUJA AND OTHER SELECTED LANGUAGES

3.7.1 Statistical data presented

3.7.1.1 Statistical data for the IR

	Languages	IR	% difference from preceding pair
1	Unguja+Mvita	2089	-
2	Unguja+Giryama	612	70.7
3	Unguja+Saghala	207	66.2
4	Unguja+Kikuyu	198	4.3
5	Unguja+Dawida	171	13.6

Table : Hierarchy of closeness of relationship between Unguja and other Selected Languages according to the IR.

3.7.2 Individual statistical characteristics

3.7.2.1 Shared reflexes

	Languages	Shared reflexes	Difference from preceding pair
1	Unguja+Mvita	507	-
2	Unguja+Giryama	348	159
3	Unguja+Kikuyu	205	143
4	Unguja+Saghala	181	24
5	Unguja+Dawida	165	16

Table : Degrees of difference in the number of shared reflexes among pairs of Selected Languages involving Unguja.

3.7.2.2 Unique reflexes

	Languages	Unique to 1st	Unique to 2nd	Total unique	Difference from preceding pair
1	Unguja+Mvita	158	56	214	-
2	Unguja+Giryama	324	112	436	222
3	Unguja+Saghala	487	114	601	165
4	Unguja+Dawida	506	133	639	38
5	Unguja+Kikuyu	458	205	663	24

Table : Degrees of difference in the number of unique reflexes among pairs of Selected Languages involving Unguja.

3.7.3 Summary of statistical results interpreted

a) Unguja and Mvita have the highest 'simple' IR at 2089. All investigated evidence suggests the overall highest figure, viewed from the sides of both languages. This is an exceptionally straightforward relationship - indeed the only twofold 'simple' one encountered in this study.

b) The lowest IR is that of Unguja and Dawida, which is one of the two overall smallest figures obtained throughout this study: 171, the overall smallest figure being 170.

It is of some interest that this IR is not in agreement with the 'proximity factor' condition.

The low IR is a consequence of a substantially low number of shared reflexes in particular, but also of a particularly high number of reflexes unique to Unguja (the overall highest at 506). As for Kikuyu, it has the unusual quality of decreasing the number of its unique reflexes down the hierarchy of closeness, reaching its low with Unguja.

c) Similarly, two basic groupings of Selected Languages are to be distinguished according to differences between their IR's.

In the first group, as in the Section on Mvita, the three 'coastal' languages would be placed, which have very high IR's with each other. According to more rigorous criteria this group would be divided further so that Mvita and Unguja would form a group of their own. As stated earlier on, they have an IR so high as to be *beyond comparison* with the remaining ones.

The second group, again in parallel with the results of the previous Section on Mvita, contains Saghala, Kikuyu and Dawida which have all comparatively low IR's with Unguja at 207, 198 and 171 respectively.

d) There is a sizable difference between the two groups of relationships outlined above: 66.2%, but differences within the second group are small. They are incomparably smaller than those in the first group. Yet Saghala and Kikuyu still have their IR's with Unguja much more similar than Kikuyu and Dawida do.

The Section on Mvita, let it be added, yielded results not dissimilar to those just described.

3.7.4 Inferences from statistical characteristics

The following points summarize the principal inferences that were drawn concerning relationships among the Selected Languages focussed on Unguja.

Certain references will be made to the preceding Section on Mvita as the two Swahili dialects display traits of common effects on relationships among the Selected Languages.

1) The closest relationship inferred in this study is indisputably that of Unguja and Mvita, with an IR of 2089. As described in the Section on Mvita, it outnumbers several-fold the next lower figure. Moreover, it is the only twofold 'simple' relationship,

as all indices on both sides of the relationship indicate.

It is highly probable that an immediate link existed between them in prehistory, and was severed rather 'recently' in the PB-B or Bantu period.

2) Unguja and Giryama are second in order of closeness as are Mvita and Giryama in the preceding Section. Nonetheless, the relationship is considerably remoter than that of Mvita and Unguja. The difference between the first two pairs is the greatest obtained in this study: 70.7%.

But 612 is a high enough IR for Unguja and Giryama to be classed as closely related, and for an immediate link in prehistory between them to be postulated. A split between them probably occurred at a considerably earlier stage in PB-B than between the first two languages.

3) The third pair, Unguja and Saghala, are likely to have separated at a very early stage in PB-B and, according to their rather low IR of 207, and an immediate link between them in prehistory is not postulated.

Unguja and Saghala's rather remote relationship represents a break from the previous two close ones, and is followed by another two which are remoter still.

The situation just described corresponds closely with the inferences made concerning Mvita and other Selected Languages.

4) The last two languages differ only to a small degree in the relationships they enjoy with Unguja. Their IR's are quite close to the preceding one, with differences of 4.3 and 13.6%.

The close quantitative resemblance of their relationships with Unguja serves as a basis for grouping Saghala, Kikuyu, Dawida

together, as opposed to the grouping together of Mvita and Giryama, which have close relationships with Unguja, No immediate link in prehistory is therefore inferred either between Unguja and Kikuyu, IR=198, or Unguja and Dawida, IR=171. Of the two pairs, the latter is somewhat isolated from the preceding two, which in turn implies that the earliest split occurred between Unguja and Dawida, followed by that between Unguja and Kikuyu.

3.7.5 Conclusions

The inferences made from statistical evidence concerning Ungujaas a focus of relationships correspond broadly with those made concerning Mvita.

Firstly, there is the evidently closest relationship of Unguja and Mvita, which is close beyond comparison with the other pairs. It can be inferred that their immediate link in prehistory disintegrated relatively 'very recently' in the PB-B or the Bantu period, while all the other pairs enjoyed separate developments from a much earlier date in prehistory.

However, the points on the relative time-scale are not easy to determine at which individual fissions, or clusters of fissions, occurred.

The evidence is convincing enough to indicate that preceding the split between Unguja and Mvita occurred that between Unguja and Giryama, as implied from the latter's high IR (612), and supported by unambiguous evidence.

Saghala's position as third in the hierarchy is borne out by explicit evidence, too. Its relationship with Unguja is rather remote. No immediate link in prehistory is postulated , although,

at a 'much earlier' stage in PB-B they did share a common ancestor together with the last two languages in the hierarchy, Kikuyu and Dawida.

Kikuyu and Unguja represent a certain anomaly in that they enjoy a closer relationship with each other than Dawida and Unguja do in spite of the fact that they are more distant geographically than the latter two languages. This third instance of the 'proximity factor' failing to operate is borne out by evidence in the Section on the relationships of Kikuyu.

To account for this phenomenon in detail would demand further study to be undertaken, and is beyond the scope of the present work. But it might be worth repeating here that Kikuyu is the only language for which the data used were not the author's own.

Dawida evidently has the remotest relationship with Unguja, with no immediate link in prehistory being likely to have existed between them.

The same grouping of languages can be outlined as was done in the Section on Mvita since conclusions concerning relationships among the Selected Languages are extremely similar in the two 'Swahili Sections'.

According to their widely different degrees of closeness of relationships that are focussed on Unguja, the languages may be grouped as follows:

1. Unguja and Mvita
2. Giryama
3. Saghala, Kikuyu, Dawida

Alternatively, the first two groups could be brought together provided that more relaxed criteria were employed.

4 SUMMARY AND CONCLUSIONS

4.1 The composite hierarchical table of closeness of relationship

The numerical results of the statistical investigations discussed in the previous chapter are presented in the composite table below.

Although the figures have already been quoted throughout chapter 3, their values in relation to the remaining pairs of languages were not immediately apparent. The figures below convey at a glance the separating or uniting tendencies among the languages in question.

In order that the significance ba made apparent of the values of the constituent figures of each pair, the numbers of unique reflexes are presented in three separate columns (unique to first language, unique to second language, total unique). The two figures, represented in the IR formula as 'x' and 'y', reveal both the internal aspects of a relationship and the eventual alternative foci of relationship.

All figures in the composite table refer to d i r e c t cognates of Common Bantu.

4	+ D	172	24	+ U	112	324	436	37	4	+ D	294	19.2
5	+ K	168	4	+ K	274	225	499	63	5	+ K	211	28.2
1	K + U	205	-	K + D	263	146	409	-	1	K + D	221	-
2	+ M	170	35	+ S	296	168	437	28	2	+ G	211	4.5
3	+ G	168	2	+ G	225	274	499	62	3	+ U	198	6.2
4	+ D	141	27	+ M	228	390	618	119	4	+ S	185	6.6
5	+ S	130	11	+ U	205	458	663	45	5	+ M	170	8.1
1	M + U	507	-	M + U	56	158	214	-	1	M + U	2089	-
2	+ G	333	174	+ G	214	128	342	128	2	+ G	737	64.7
3	+ S	180	153	+ S	385	115	500	158	3	+ S	247	66.5
4	+ K	170	10	+ D	408	142	550	50	4	+ D	177	28.3
5	+ D	151	19	+ K	390	228	618	68	5	+ K	170	4
1	U + M	507	-	U + M	158	56	214	-	1	U + M	2089	-
2	+ G	348	159	+ G	324	112	436	222	2	+ G	612	70.7
3	+ K	205	143	+ S	487	114	601	165	3	+ S	207	66.2
4	+ S	181	24	+ D	506	133	639	38	4	+ K	198	4.3
5	+ D	165	16	+ K	458	205	663	24	5	+ D	171	13.6

Composite Hierarchical Table of Closeness of Relationship.

D = Dawida G = Giryama M = Mvita
S = Saghala K = Kikuyu U = Unguja

Number of Shared Reflexes				Number of Unique Reflexes				Index of Relationship			
	Languages	Shared reflexes	Difference from preceding pair	Languages	Unique to 1st	Unique to 2nd	Total unique	Difference from		Languages	IR	% difference from preceding pair
1	D + S	188	-	D + S	105	110	215	-	1	D + S	642	-
2	+ G	172	16	+ G	116	283	399	184	2	+ G	294	54.2
3	+ U	165	7	+ K	146	263	409	10	3	+ K	221	24.8
4	+ M	151	14	+ M	142	408	550	141	4	+ M	177	19.9
5	+ K	141	10	+ U	133	506	639	89	5	+ U	171	3.4
1	S + G	196	-	S + D	110	105	215	-	1	S + D	642	-
2	+ D	188	8	+ G	109	270	379	164	2	+ G	364	43.3
3	+ U	181	7	+ K	168	269	437	58	3	+ M	247	32.1
4	+ M	180	1	+ M	115	385	500	63	4	+ U	207	16.2
5	+ K	130	50	+ U	114	487	601	101	5	+ K	185	10.6

4.2 Method of investigation

In order to investigate relationships among the six languages in some detail, all combinations of pairs of languages were determined (see the Triangle of Relationships, p. 123). Fifteen pairs were established which are in the focus of the following discussion.

Only d i r e c t C.B. cognates were used in the computations with a view to produce statistical data directly comparable with Guthrie's and further our present knowledge of Bantu languages in a direct manner.

4.2.1 Indicators of closeness of relationship

The key indicator of closeness of relationship adopted in this study is the Index of Relationship, IR, calculated on the basis of Guthrie's formula.

The results arrived at by the shared reflex index and the simple index, SI, both of which are parts of the IR formula, do not always coincide with the IR results. The shared reflex index on its own certainly appears insufficient as an indicator of closeness for purposes of this study.

In this study only nine relationships out of thirty can be termed 'simple', i.e. formed by indices that are hierarchically identical in relation to other pairs of a particular group.

4.2.2 Group analyses

Six groups of relationships were studied, each of which was focussed on one Selected Language and consisted of five pairs.

Thus each relationship was investigated from two sides, producing, needless to say, not necessarily identical results. For example, the

'simple' close relationship of Dawida and Saghala, observed from Dawida's side, yielded a different picture when considered as the Saghala+Dawida pair, observed from Saghala's side. As such it could not be termed 'simple' because one of the indices disturbed the place of order determined by the other two.

Each group produced results different from the others. However, certain groups tended to coincide in broad terms of the relationships they formed within themselves, thus indicating certain general tendencies.

The two Swahili dialects are at the centre of two groups which produced results of substantial similarity. Their resulting hierarchies as well as the internal characteristics of those hierarchies are very much alike. Unguja and Mvita with the closest relationship are followed by Giryama and Saghala in both groups in that order, but positions of Dawida and Kikuyu are reversed.

Their internal characteristics reveal a large difference between Mvita, Unguja and Giryama with a great degree of closeness among them on one side and the remaining languages, with smaller degrees of closeness between each other, on the other. A further division would affect the first group and separate Giryama from Mvita and Unguja the relationship of which is close beyond comparison with the rest.

Another two groups, those centered round Giryama and Saghala, also yielded results which, in broad terms of reference, were not dissimilar from the joint outcome of the Swahili groups.

The G i r y a m a-centered group produced a hierarchy identical in rank order with the Mvita group, but with different internal proportions.

However, the division into two broad groupings, as indicated in the Swahili groups, can be applied unaltered to this group also: Giryama, Mvita, Unguja on one side, and Saghala, Dawida and Kikuyu on another.

The S a g h a l a-centered group, while setting aside the relationship

with Dawida as one specially close, underlined still further the tendency towards the grouping together of Giryama, Mvita and Unguja as the three preceding groups did. The position of Kikuyu in this group indicates its isolation from the remaining languages.

Observed from the angle of Dawida, the picture is somewhat changed, but the obvious closeness of Mvita and Unguja is distinctly borne out. The other apparent grouping is here formed by Dawida and Saghala, which have an essentially close relationship.

Kikuyu, as a focus of relationships, appears to be the one exception among the six groups. The hierarchy it effects is not only at variance with the other groups, but it also contains such small and even differences within the IR as indicate no recognizable groupings. In the other hierarhcies Kikuyu occupies the last position except in the Dawida and Unguja groups.

Further evidence of the special position of Kikuyu among the Selected Languages is offered by structural data. Kikuyu is the only seven-vowel language, it displays exceptionally straightforward sound-correspondences and has an unusually small number of skewed reflexes. It is the only language among the six with a single explicit reflex of *d, and its consonantal correspondences with C.B. are identical in open-and close-vowel environment.

4.3 Results of investigation

On the basis of the discussed hierarchies of closeness of relationship certain trends are to be discerned towards groupings of the languages according to the measure of closeness.

The following tentative demarcations can be drawn:

(I)

IA Mvita, Unguja

IB Giryama

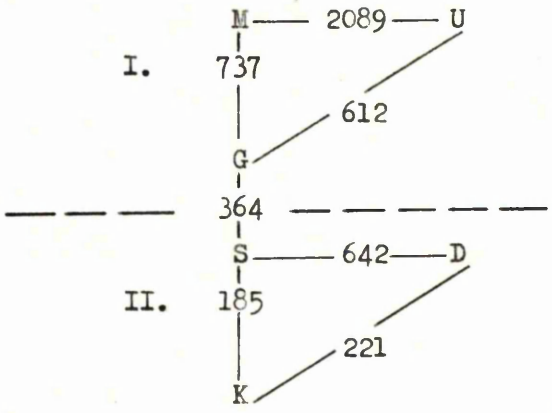
(II)

IIA Saghala, Dawida

IIB Kikuyu

These demarcations are of necessity gross generalizations and as such they might be subject to misinterpretation. It is essential to realize that the degrees of relationship within and between the groups are far from identical.

The following diagram illustrates the actual dimensions expressed by IR figures. Distances between letters do not correspond to either linguistic or geographical proximity of the languages they represent.



Index of Relationship proportions within and between the two basic groupings of Selected Languages.

According to the above diagram,

1) relationships in group I are generally much closer than those in group II,[1]

2) group I contains a relationship that is much closer than any other,

3) IIB, Kikuyu, has relatively remote relationships with all members of both groups.

The grouping just described was corroborated by results of several computer programmes which used the IR and SI figures as well as the shared reflex figures.[2]

In the first programme a classification was worked out according to the IR, SI and the shared reflex indices. The IR and SI imply the same hierarchic relationships. The shared reflex figures alone imply a slightly different hierarchy, but even there the most reliable strategy (group average classification)[3] only differs from the results of other classifications in the relative closeness it implies between Mvita+Unguja+Giryama and Kikuyu and Saghala+Dawida respectively.

The second programme established various groupings of the languages by the method of overlapping clusters. In all cases Kikuyu stood out as the 'single object'[4] against the remaining languages, Mvita+Unguja formed by all accounts the closest group. To demonstrate the results of the programme, the groupings established in its second part (groups overlap

---

1 The adverb 'generally' indicates that there is a group II relationship which is closer than at least one group I relationship. It is Dawida+Saghala in respect of Giryama+Unguja.

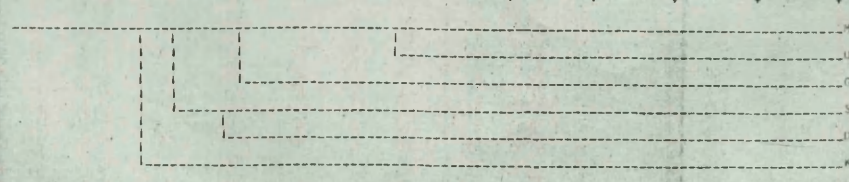
2 I am thankful to M. Mann who prepared the programmes. The methods used in the programmes have been described by A. Henrici. See A. Henrici, op. cit.

3 Information supplied by M. Mann.

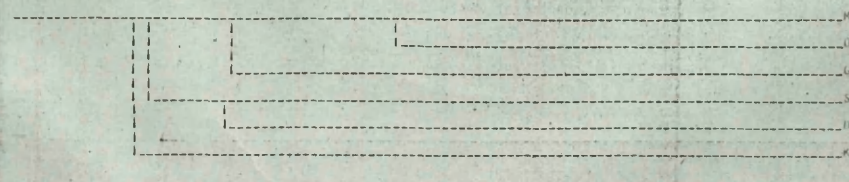
4 Computer programme term.

SIX NORTH-EAST BANTU LANGUAGES: CLASSIFICATION BY GUTHRIE'S IR

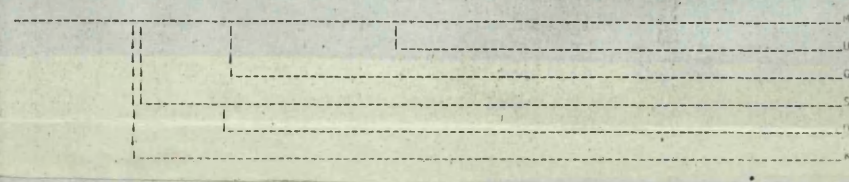
SIX NORTH-EAST BANTU LANGUAGES: CLASSIFICATION BY GUTHRIE'S IR — NEAREST NEIGHBOUR CLASSIFICATION



SIX NORTH-EAST BANTU LANGUAGES: CLASSIFICATION BY GUTHRIE'S IR — BRANCH AVERAGE CLASSIFICATION

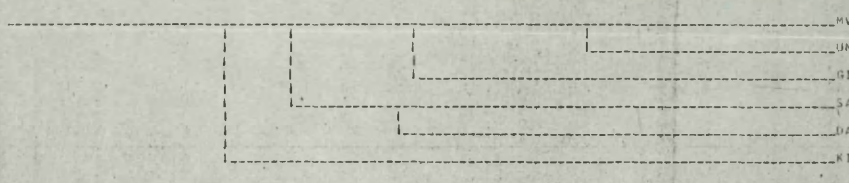


SIX NORTH-EAST BANTU LANGUAGES: CLASSIFICATION BY GUTHRIE'S IR — GROUP AVERAGE CLASSIFICATION

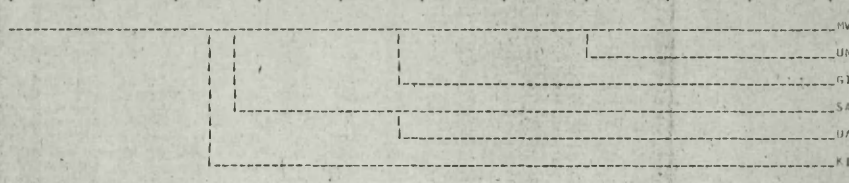


SIX NORTH-EAST BANTU LANGUAGES: CLASSIFICATION BY SIMPLIFIED IR (SI)

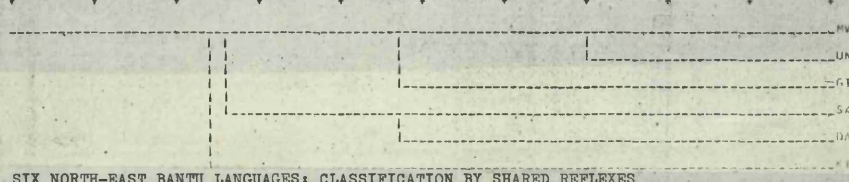
SIX NORTH-EAST BANTU LANGUAGES: CLASSIFICATION BY SIMPLIFIED IR — NEAREST NEIGHBOUR CLASSIFICATION



SIX NORTH-EAST BANTU LANGUAGES: CLASSIFICATION BY SIMPLIFIED IR — BRANCH AVERAGE CLASSIFICATION

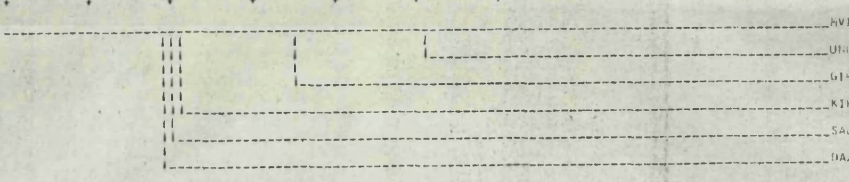


SIX NORTH-EAST BANTU LANGUAGES: CLASSIFICATION BY SIMPLIFIED IR — GROUP AVERAGE CLASSIFICATION

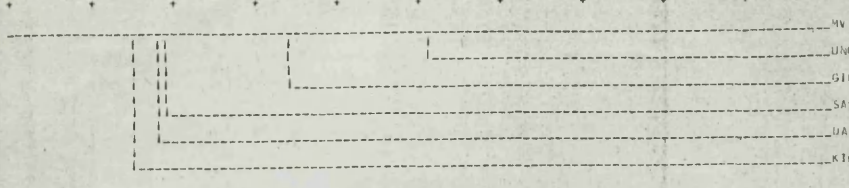


SIX NORTH-EAST BANTU LANGUAGES: CLASSIFICATION BY SHARED REFLEXES

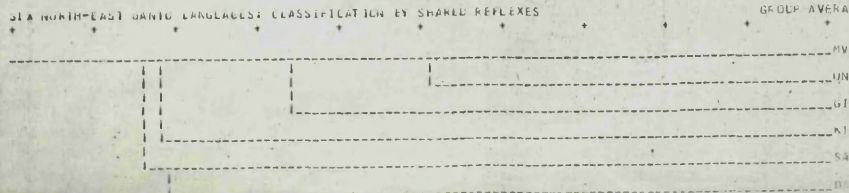
SIX NORTH-EAST BANTU LANGUAGES: CLASSIFICATION BY SHARED REFLEXES — NEAREST NEIGHBOUR CLASSIFICATION



SIX NORTH-EAST BANTU LANGUAGES: CLASSIFICATION BY SHARED REFLEXES — BRANCH AVERAGE CLASSIFICATION

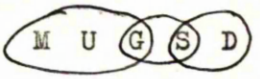
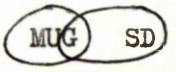


SIX NORTH-EAST BANTU LANGUAGES: CLASSIFICATION BY SHARED REFLEXES — GROUP AVERAGE CLASSIFICATION



by one language) will be presented below. The groupings in the left-hand column will be graphically represented on the right.

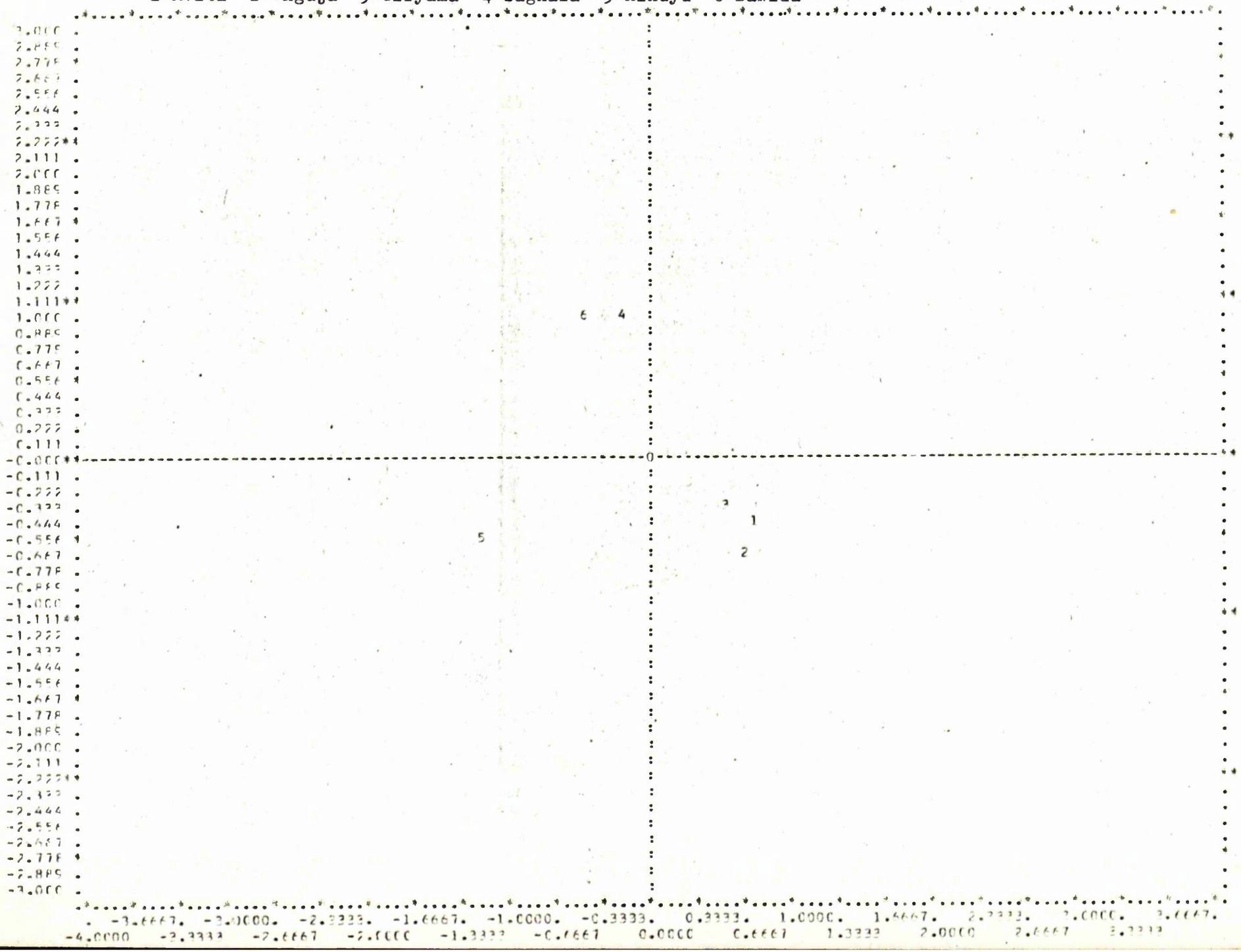
Selected Languages: grouping by overlapping clusters based on IR figures.
(Groups overlap by one language.)
M=Mvita, U=Unguja, G=Giryama, S=Saghala, D=Dawida, K=Kikuyu.

Cluster level 1	M+U	
Cluster level 2	M+U M+G	
Cluster level 3	M+U M+G S+D	
Cluster level 4	M+U+G S+D	
Cluster level 5	M+U+G G+S S+D	
Cluster level 6	M+U+G G+S+D	
Cluster level 7	M+U+G+S+D	
Cluster level 8	M+U+G+S+D K+D	
Cluster level 9	M+U+G+S+D+K	

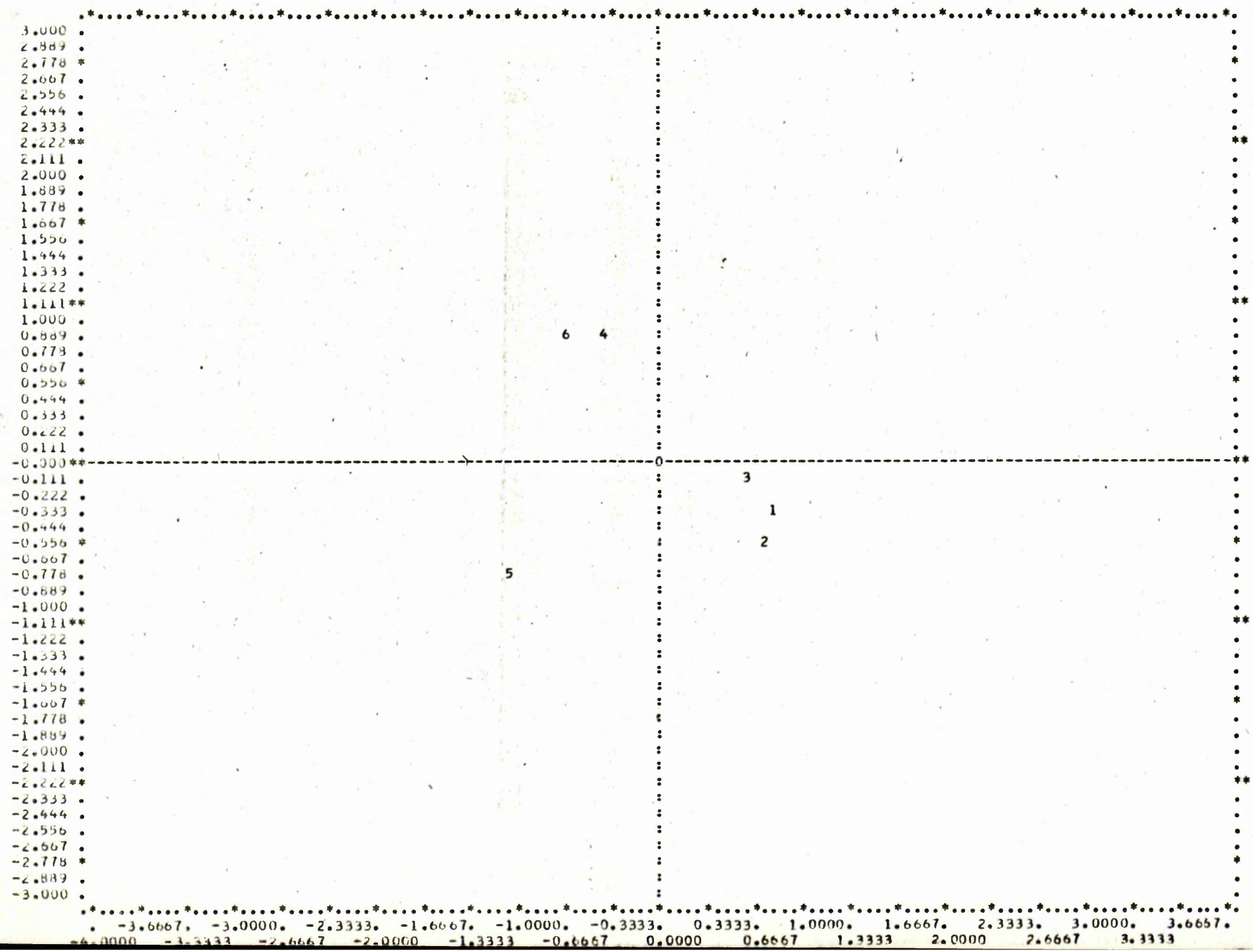
The third programme produced two schematic maps of the Selected Languages, also on the basis of the IR and SI figures. Note that it only seeks to place languages relatively nearer or further apart in space where they are linguistically more or less similar; it does not claim to place them proportionately nearer or further, i.e. on a strict scale, according to the degree of linguistic similarity.

TWO SPACE CONFIGURATION FOR MAPPING OF SIX NORTH-EAST BANTU LANGUAGES ON THE BASIS OF GUTHRIE'S 18

1=Mvita 2=Unguja 3=Giryama 4=Saghala 5=Kikuyu 6=Dawida



1=Mvita 2=Unguja 3=Giryama 4=Saghala 5=Kikuyu 6=Dawida



4.3.1 Summary of closeness of relationship

It can be *inferred*, according to the key indicator of degrees of the IR,

1 that M+U are the closest related languages in the group,

2 that the next closest relationship is that of G+M,

3 that D+S are less closely related than G+M, but more closely than all the remaining pairs,

4 that S+G have a closer relationship than D+K do,

5 that S+K have a remoter relationship than D+K, and much remoter than S+M,

6 that S+M are more closely related than D+K,

7 that K has a closer relationship with U than with M,

8 that D has a closer relationship with M than with U,

9 that S has a closer relationship with M than with U, and more so than does D.

4.4 Diachronic inferences

The final product of the statistical examination is, in its most generalized form, a division of the Selected Languages into groups, the relationships within and between which may be interpreted diachronically.

The interpretation put forward here is strictly tentative and should ideally be approached only with reference to the actual statistical and structural data.

When put onto the relative time-scale the two basic groups of relationships (see p. 178) would appear to have split at different points. But the fission of Dawida and Saghala, group IIA, may be implied as having occurred at approximately the same time on the relative time-scale as that of the first group, IA and IB, according to the values of their IR's.

There is also the specially remote position of Kikuyu, which is not apparent from the simple grouping of I and II. Kikuyu, according to its lowest IR figures with three languages, and low with two, imply the least close relationship and the earliest fission from the group of six.

4.4.1 A tentative genealogy of the Selected Languages

Having accepted the hypothesis of a common origin

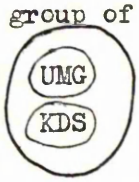
1) for the Bantu languages as a whole,
2) for languages classified broadly as belonging to the Eastern of the Bantu field, we can consider a hypothesis of developments that effected the origins of the individual Selected Languages.

The tentative genealogy propounded below was formed on the basis of both the actual detailed data and on the generalized results that further assist in outlining the broader tendencies among the languages.

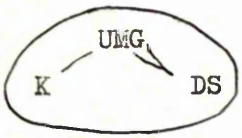
It would appear, on the evidence available, that Kikuyu separated

first from the remaining languages and the common ancestor they had shared. To illustrate the whole postulated process, it will be summarized in a number of progressive steps.

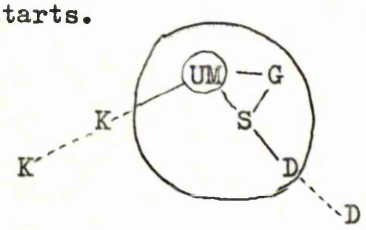
1 Polarization within the group of languages in PB-B starts



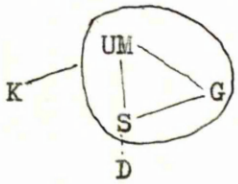
2 Polarization progresses.



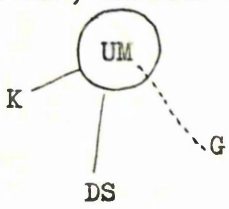
3 Extreme polarization. The fission of K in progress, fission of D starts.



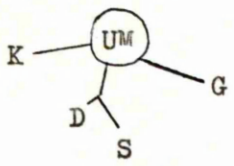
4 Fission of K completed, progressive fission of D+S starts. Simultaneously, G starts to polarize towards fission.



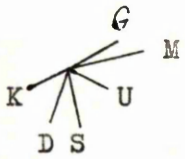
5 Fission of D+S completed, fission of G starts.



6 D and S split, fission of G completed.



7 U and M split.



4.5 A note on Taita: in conclusion

We have attempted to answer the question which this study posed itself, namely what relative degrees of relationship there are between the six Selected Languages. At the same time the more specific question was being studied, which stood at the outset of this study, of the hitherto unresearched relationship of Dawida and Saghala.

While the allegedly close link between Saghala and Giryama has been confirmed by the outcome of this study, a similarly close link between Dawida and Kikuyu proved not to be of any substantial strength although it had been considered a strong possibility by some Bantuists, and on isolated pieces of evidence.

In fact Dawida has been demonstrated to have a somewhat closer relationship with Giryama than with Kikuyu.

While there is no doubt that the single unit classification of the two Taita languages has been justified, their status as dialects, as opposed to separate languages, certainly has not.

Compared with the close relationships between Mvita and Unguja, or Zulu and Xhosa, which have an even closer relationship with each other than the two Swahili dialects do, theirs is a considerably less close one.

It is hoped that, on the basis of the outcome of

this investigation, further questions will be raised concerning the relationships between languages in the north-eastern area of the Bantu field.

BIBLIOGRAPHY

(Only those works are listed which were directly used in the thesis.)

Patrick R. Bennett, 'Identification, classification, and Bantu linguistics', ALS, XIV, 1973, 19-25.

T. G. Benson, (ed.) Kikuyu-English dictionary, Oxford at the Clarendon Press, 1964.

Rev. Canon Binns, Swahili-English dictionary, London, S.P.C.K., 1925.

P.G.Bostock, The peoples of Kenya: The Taita, MacMillan, London, 1950.

M.A.Bryan, Handbook of African languages: The Bantu languages of Africa, O.U.P. for International African Institute, 1959.

Theodora Bynon and Michael Mann, 'An introduction', ALS, XIV, 1973, 1-5.

Hazel Carter, 'Tonal data in "Comparative Bantu", ALS, XIV, 1973, 36-52.

A.M.Champion, The Agiryama of Kenya, (ed. J.Middleton), Occasional paper No.25, Royal Anthropological Institute, London, 1967.

D.T.Cole, 'Doke's classification of the Bantu languages', African Studies,XXX, 3/4, 1971.

Florence Deed, Giryama-English dictionary, East African Literature Bureau, 1964.

Florence I.Deed, Giryama exercises, [Cyclostyled undated copy.]

C.M.Doke, Bantu: Modern grammatical, phonetical, and lexicographical studies since 1860, Dawsons of Pall Mall, London, for International African Institute, 1967.

Joseph H.Greenberg, The languages of Africa, Bloomington, 1963.

Joseph H.Greenberg, 'Linguistic evidence regarding Bantu origins', JAH, XIII, 2, 1972, 189-216.

A.Hémery, Vocabulaire francais-swahili-teita, Zanzibar and Paris, 1901.

Alick Henrici, 'Numerical classification of Bantu languages', ALS, XIV, 1973, 82-104.

H.H.Johnston, The Kilima-njaro expedition [...], London, 1886.

Rosalie Jones-Phillipson, Affinities between Venda and other southern Bantu languages, Ph.D. thesis, University of London, 1972.

L.Krapf, A dictionary of the Suahili language, London, Trübner & Co., Ludgate Hill, 1882.

L.Krapf, Outline of the elements of the Kisuaheli language, with special reference to the Kinika dialect, Tübingen, 1850.

J.L.Krapf, Vocabulary of six East-African languages (Kisuaheli, Kinika, Kikamba, Kipokomo, Kihiau, Kigalla.) Tübingen, 1850.

J.L.Krapf and J.Rebmann, A Nika-English dictionary, S.P.C.K., London, 1887.

W.Michael Mann, 'Internal relationships of the Bantu languages: prospects for topological research', in Language and History in Africa, ed. David Dalby, London, 1970.

Michael Mann, 'Sound-correspondences and sound-shifts', ALS, XIV, 1973, 26-35.

Archdeacon Maynard, An English-Kidawida dictionary, Copied from Ms. in possession of Miss M.H.Murray. December 1950. This copy has a note at the end 'Finished Feb.21. 07'.

Archdeacon Maynard, Kidawida grammar, preserved by Miss Murray in 1920. [Typewritten] Copy by Education Office, Wundanyi, Teita District.

A.E.Meeussen, 'Test cases for method', ALS, XIV, 1973, 6-18.

D.Nurse and G.Philippson, The north-eastern Bantu languages of Tanzania and Kenya: a classification, The Institute of Swahili Research, University of Dar es Salaam, 1974.

D.J.Parkin, Palms, wine and witnesses, Intertext Books, London, 1972.

Charles Sacleux, Dictionnaire swahili-francais, Paris, Institut d'Ethnologie, Musée de l'Homme, 1939.

C.Sacleux, Grammaire des dialectes swahilis, Paris, 1909.

A.Downes Shaw, A pocket vocabulary of the Ki-swahili, Ki-nyika, Ki-taita, and Ki-kamba languages, London, 1885.

Magdalena Slavikova and Margaret M.Bryan, 'Comparative Bantu: the case of two Swahili dialects', ALS, XIV, 1973, 53-81.

C.H.Stigand, A grammar of dialectic changes in the Kiswahili language, Cambridge at the University Press, 1915.

Harry R.Tate,'Short vocabulary of Teita', in Journal of the Royal Anthropological Institute, 34, 1904, 140-48.

W.E.Taylor, Giryama vocabulary and collections, S.P.C.K., London, 1891.

A.N.Tucker and M.A.Bryan, Linguistic survey of northern Bantu border land, 1, III. Far Eastern Section. Great Lakes to Indian Ocean, O.U.P. for International African Institute, 1956.

A.N.Tucker and M.A.Bryan, Linguistic survey of the northern Bantu borderland, Vol.4, O.U.P. for International African Institute, 1957.

E.O.J.Westphal, 'On classifying Bushman and Hottentot Languages', ALS,III, 1962, 30-48.

J.Williamson,'Dabida numerals', African Studies, II, 4, 1943, 215-6.

H.W.Woodward, 'Kitaita or Kisighau, as spoken on the Shambala hills above Bwiti', Zeitschrift für Kolonialsprachen, IV, 2, 91-117.

J.Alfred Wray, An elementary introduction to the Taita language, Eastern Equatorial Africa, S.P.C.K., London, 1894.

WORKS BY MALCOLM GUTHRIE

which were directly used in the work.

'Bantu origins: a tentative new hypothesis', JAL, I, 1, 1962, 9-21.

'Some developments in the prehistory of Bantu languages', JAH, III,2, 1962, 273-82.

'A two-stage method of comparative Bantu study', ALS, III,1962,1-24.

'Some uses of arithmetical computation in comparative Bantu studies', Transactions of the Philological Society 1964, 1965,108-28.

'Comparative Bantu: a preview', JAL,IV,1,1965,40-45.

Comparative Bantu: an introduction to the comparative linguistics of the Bantu languages, 4 volumes, Gregg Press Ltd., Farnborough, Hants, England, 1967-1971.

'Contributions from comparative Bantu to the study of African prehistory', in Language and History in Africa: collected papers of the London seminar on language and prehistory in Africa, Frank Cass, London, 1971.

Collected papers on Bantu linguistics, Gregg International Publishers Limited, England, 1970.

Note: For a complete list of Guthrie's writings see Guy Atkins, 'Writings of Malcolm Guthrie', ALS,XI,1970,2-4.